SOX2: Functional analysis using mouse mutants

Thesis

UNRESTRICTED.

SOX2: Functional analysis using mouse mutants.

Donald McIntyre Bell BSc Hons. (Biochem.) Edinburgh.

August 2004

Thesis submitted for the degree of Doctor of Philosophy with

The Open University

Sponsoring establishment

Department of Developmental Genetics
The National Institute For Medical Research
The Ridgeway
Mill Hill
London
NW7 1AA

Submission date: 3 August 2004
Award date: 7 March 2005

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

Index

Index of figures

Index of Tables

Abbreviations

ßgal	betagalactosidase
ßgeo	betagalactosidase/neomycin fusion gene
bp	base pairs
cm	centimetre
CM	centimorgan
CNS	central nervous system
ddH_2O	double distilled water
°C	degrees Celsius
DNA	deoxyribonucleic acid
dpc	days post coitum
ES cells	embryonic stem cells
EST	expressed sequence tag
IRS	inner root sheath
kb	kilobase
M	molar
H3P	phosphorylated histone H3
neo	neomycin
ORS	outer root sheath
PBS	phosphate Buffered Saline
PBS^{CMF}	phosphate Buffered Saline (calcium and magnesium free)
PCR	polymerase chain reaction
PGK	phosphoglycerate kinase
PNK	polynucleotide kinase
PNS	peripheral nervous system
RMS	rostral migratory stream
RNA	ribonucleic acid
SSC	standard Saline Citrate
SDS	sodium dodecyl sulphate
UTR	untranslated region
UV	ultraviolet
w/v	weight to volume
X-gal	5-Bromo-4-chloro-3-indolyl beta-D-galactopyranoside

Abstract

The *Sox* genes are a large group of transcription factors with representatives in all animal species and initial investigations have demonstrated the importance of many members of this family during embryogenesis.

The expression pattern of *Sox2* was surveyed in the second half of gestation by examining the X-gal staining pattern of $Sox2^{\beta geo-/+}$ embryos in wholemount and section immunohistochemistry using a SOX2 specific polyclonal antibody. Particular attention was given to the pattern in the developing inner ear and in the hair follicle. Expression suggests that *Sox2* plays significant roles in a diverse range of tissues throughout embryogenesis.

The null mutation of *Sox2* demonstrates a critical requirement for this gene in embryogenesis but early lethality of $Sox2^{\beta geo-/-}$ embryos precludes study at later stages. A strategy to produce a conditionally null allele of *Sox2* was devised to circumvent the early lethality and allow functional investigation of *Sox2* in a tissue specific and/or time dependent manner throughout embryogenesis and post-natally. Initial attempts to target the *Sox2* allele were unsuccessful but redesigning the targeting vector resulted in the production of correctly targeted ES cells. Targeted clones were used to produce high percentage chimeras but the ES cell component did not contribute to subsequent generations.

Two mouse mutants, *Lcc* and *Ysb*, display abnormalities in inner ear and coat formation. Expression analysis and physical mapping implicate *Sox2* as a strong candidate affected by these alleles. The creation of mice carrying combinations of $Sox2^{\beta geo}$, *Lcc* and *Ysb* mutations showed that they are allelic and that inner ear and coat abnormalities are caused predominantly by the loss of expression of *Sox2*. Congenic breeding of $Sox2^{\beta geo}$

mice onto a C3H/He background revealed heterozygous coat phenotypes confirming an important role for *Sox2* in hair formation.

The closely related gene *Sox1* is co-expressed along with *Sox2* during CNS development and physical similarities between their protein products suggest overlapping functions able to compensate for the loss of either gene. Breeding experiments demonstrate an increase in the rate of mortality far greater than that expected by the addition of individual phenotypes. Immunohistochemical analysis reveals aberrant cell behaviour in CNS progenitor cell populations carrying the most severe genotype and suggests some common function between these genes during CNS development.

1.2 Sox Proteins and genes - Classification and phylogeny

The identification of members of the Sox gene family grew from the discovery of the mammalian sex determining factor SRY reported in 1990 (Gubbay et al., 1990; Sinclair et al., 1990). SRY, which is found on the Y chromosome in mice and humans, has a DNA binding domain similar to the non-histone chromosomal proteins HMG-1 and HMG-2. The HMG domain or box can be found in a large number of genes although few positions in the 70-80 amino acid long region are conserved throughout the superfamily (Laudet et al., 1993)). The broadest division of the HMG-box proteins separates members of the HMG/UBF families from those of the SOX/TCF/MATA genes (Soullier et al., 1999). HMG/UBF genes are found in the three major eukaryotic phyla (fungi, plants and animals) and generally contain multiple DNA binding domains that associate with DNA with low to moderate affinity and little sequence specificity (Soullier et al., 1999). The SOX/TCF/MATA family usually have a single HMG domain that binds specifically to target sequences. This family appears to be younger, with MATA and SOX/TCF genes found exclusively in yeast and animals respectively. Significant diversification is restricted to the SOX/TCF genes in animals with the SOX family having more than thirty members. The human and mouse genomes each contain twenty SOX genes with further genome duplication thought to be responsible for additional family members found in other vertebrates such as zebrafish and Xenopus (Bowles et al., 2000; Pevny and Lovell-Badge, 1997; Schepers et al., 2002; Soullier et al., 1999; Wegner, 1999).

Conventionally, as the family name indicates, the *SOX* gene family is based

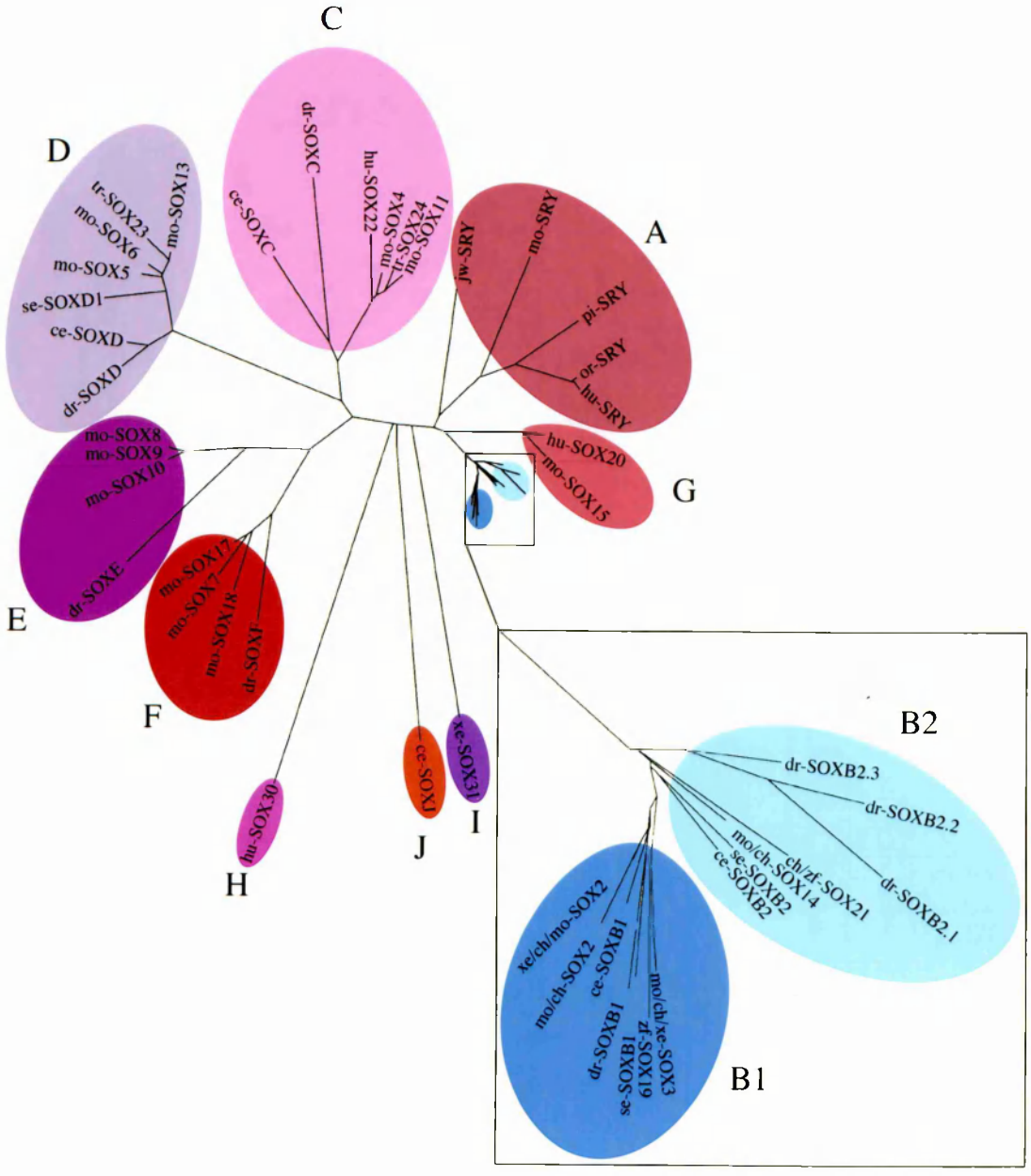
on the criterion of at least 50% identity to *Sry* within the HMG box (Collignon et al., 1996). However, homology with Sry, a gene which is exclusive to mammals, is not necessarily the most accurate method of classifying the whole Sox family. Recent analysis, benefiting from the complete HMG domain sequence of over seventy *SOX* genes isolated from nematodes, insects, amphibians, reptiles, fish, birds, and mammals, suggests an improved method of classification based on the conservation of particular HMG domain motifs (Bowles et al., 2000). This includes genes such as human-SOX30 and *Caenorhabditis elegans*-SOXJ, which have approximately 45% identity in their HMG box to *Sry* but would exclude the Zebrafish gene *casanova* (Dickmeis et al., 2001; Kikuchi et al., 2001; Sakaguchi et al., 2001). *Casanova* is one of the few family members that does not contain the exact protein motif RPMNAFMVW in its HMG box although substitutions are conservative (Dickmeis et al., 2001). This amino acid sequence represents those residues that confer specificity by having direct base contact as well as those that intercalate with the minor groove of the DNA helix facilitating architectural changes (Dailey and Basilico, 2001). It is clear that, as further related genes are discovered, only a more flexible definition will include all the Sox family members.

1.3 Sox gene subgrouping

Sequence data from many *Sox* family members in diverse species also helps to confirm their sub-grouping into categories of greatest similarly. In 1993, by comparison of the HMG domain amino acid sequence of the fifteen mouse *Sox* genes identified, members with more than eighty percent amino acid identity were grouped together, resulting in seven sub-divisions of the *Sox* gene family

Figure 1.1 An unrooted phylogeny for the SOX HMG domain.

This diagram (adapted from Bowles 2000) demonstrates the clustering of Sox *genes from diverse species on the basis of similarities in the HMG domain. Branch lengths represent the extent of divergence and colours highlight membership of a particular subgroup. A single mammalian representative of each gene is included where numerous orthologues exist (except group A). The close conservation amongst group B genes requires a magnified view. ce, nematode, Caenorhabditis elegans; dr, fruit-fly, Drosophila melanogaster; hu, human, Homo sapiens; mo or m, mouse, Mus musculus; or, orangutan, Pongo pygmaeus; pi, pig, Sus scrofa ; tw, tammar wallaby (marsupial), Macropus eugenii; tr, rainbow trout, Oncorhynchus mykiss; xe, frog, Xenopus laevis; zf, zebra fish, Danio rerio.*


C
dr-SOXC
hu-SOX22
mo-SOX4
tr-SOX24
mo-SOX11
ce-SOXC
D
tr-SOX23
mo-SOX13
mo-SOX6
mo-SOX5
se-SOXD1
ce-SOXD
dr-SOXD
A
mo-SRY
pa-SRY
pi-SRY
or-SRY
hu-SRY
G
hu-SOX20
mo-SOX15
E
mo-SOX8
mo-SOX9
mo-SOX10
dr-SOXE
F
mo-SOX17
mo-SOX7
mo-SOX18
dr-SOXF
H
hu-SOX30
J
ce-SOX1
I
xe-SOX31
B2
dr-SOXB2.3
dr-SOXB2.2
mo/ch-SOX14
ch/zf-SOX21
dr-SOXB2.1
se-SOXB2
ce-SOXB2
B1
xe/ch/mo-SOX2
mo/ch-SOX2
ce-SOXB1
dr-SOXB1
se-SOXB1
mo/ch/xe-SOX3
zf-SOX19

(Wright et al., 1993). This classification was subsequently extended as further genes were discovered leading Bowles et al to propose that there should be ten sub-groups (A-J). Three of these groups have only a single member at present with orthologues perhaps awaiting discovery (Fig. 1.1).

Except for group B, which may show a relatively recent divergence, invertebrate *SOX* genes have a single representative for each sub-group and phylogenetic analysis indicates that the numerous vertebrate subgroup members are the products of duplication events (Cremazy et al., 2001). Five Sox group B genes have been identified in vertebrates and the gene pairs SOX1/SOX21 and SOX2/SOX14 show conserved synteny in chicken and human indicating that duplications that gave rise to these genes pre-date the divergence of birds and mammals 310 million years ago (Kirby et al., 2002; Kuroiwa et al., 2002). It is therefore possible that the seven or so invertebrate Sox genes were duplicated in two steps perhaps corresponding to the proposed genome duplications that gave rise to the four HOX gene clusters in early vertebrate development (Amores et al., 1998; Kirby et al., 2002).

Amongst the vertebrate genes discovered, comparison of non-HMG domains within sub-groups generally upholds current classification even though sequence identity is much lower (Fig. 1.2). Functional studies have also revealed similarities between the closest relatives within species but although grossly these proteins are extremely similar more subtle differences may account for their conservation over such a long period of evolution (Kamachi et al., 1998; Kuhlbrodt et al., 1998a). It is tempting to speculate that greater complexity in vertebrates can, in part, be attributed

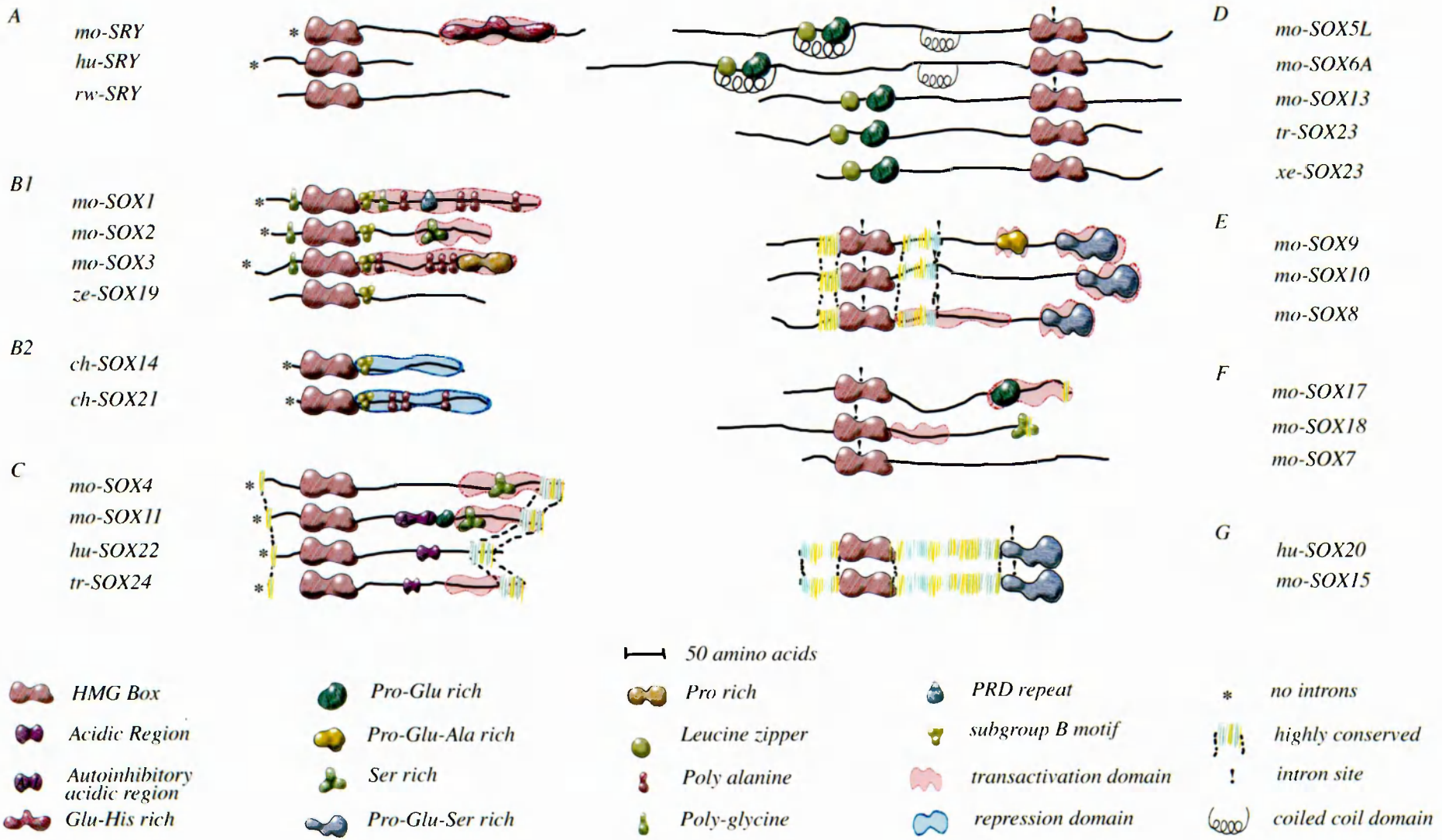

A
mo-SRY
hu-SRY
rw-SRY
B1
mo-SOX1
mo-SOX2
mo-SOX3
ze-SOX19
B2
ch-SOX14
ch-SOX21
C
mo-SOX4
mo-SOX11
hu-SOX22
tr-SOX24
D
mo-SOX5L
mo-SOX6A
mo-SOX13
tr-SOX23
xe-SOX23
E
mo-SOX9
mo-SOX10
mo-SOX8
F
mo-SOX17
mo-SOX18
mo-SOX7
G
hu-SOX20
mo-SOX15
50 amino acids
HMG Box
Acidic Region
Autoinhibitory acidic region
Glu-His rich
Pro-Glu rich
Pro-Glu-Ala rich
Ser rich
Pro-Glu-Ser rich
Pro rich
Leucine zipper
Poly alanine
Poly-glycine
PRD repeat
subgroup B motif
transactivation domain
repression domain
no introns
highly conserved
intron site
coiled coil domain

Figure 1.2 Schematic representation of SOX proteins.

This diagram (adapted from Bowles 2000) highlights the gross similarities and differences between the SOX proteins when they are arranged in groups according to their HMG domain sequence. Various structural features, motifs, and functional regions (demonstrated or putative) are shown along with intron positions and sizes where known. Genomic structures are known in some cases "" indicates that an intronless structure has been reported. Abbreviations as for fig. 1.1.*

to the development of more elaborate non-HMG domains of this large family of DNA binding proteins.

1.4 *Sox* genes are transcription factors.

The HMG domain that defines the *Sox* family facilitates the association of members with DNA where their primary function appears to be to regulate gene expression (Kamachi et al., 1998; Ng et al., 1997; Wiebe et al., 2003; Yuan et al., 1995). A gene is expressed when an RNA copy of it is produced by the action of the basal transcription machinery. The activity of this complex is often controlled by numerous *cis*-regulatory elements that might be located flanking the gene or perhaps within intronic sequence. Each regulatory element allows the binding of specific transcription factors which in turn modify the activity of the transcriptional machinery. Individual regulators may effect several genes and a gene may require the coordinated action of several transcription factors to ensure appropriate expression that is usually dynamic quantitatively spatially and temporally. Interaction of these proteins produces a regulatory network that dictates the overall expression profile and thus the identity of any given cell.

1.5 The ability of the *SOX* proteins to bind to and bend DNA.

The resolution of the crystal structure of the Sox type HMG domain reveals that the amino acids most highly conserved across the group contribute to a twisted L-shape DNA binding region and a hydrophobic core that maintains its shape. Target DNA lies across the concave surface of the 'L' that partially intercalates into the minor groove inducing a bend in the DNA of between 70° and 85°. Sox proteins show some capacity to bind particular DNA conformations without sequence

specificity *in vitro* but this ability has not yet been demonstrated *in vivo* and therefore may not be functionally important (Travers, 2000; Wegner, 1999). Instead, the function of Sox factors appears to be dependent upon a specific relationship with variants of the core consensus motif $^{A}/_{T}$ $^{A}/_{T}$ CAA $^{A}/_{T}$ G (Mertin et al., 1999).

Conformational change upon binding of a specific target sequence is a characteristic of Sox proteins and has been hypothesised to play a significant role in gene regulation by this family (Pevny and Lovell-Badge, 1997). When the group B proteins, SOX1 and SOX2, were compared to SOX9, a Group E protein, *in vitro* binding and bending assays gave almost indistinguishable results (Kamachi et al., 1999).

1.6 Transcriptional regulation by the Sox proteins.

The specific association of Sox proteins with DNA suggests a role in gene regulation and several family members display an activation capacity that is independent of cell type or SOX DNA binding domain. Discrete regions from the C terminus of several Sox proteins have been shown to possess an ability to activate reporter genes *in vitro* (Sox2 (Kamachi et al., 1998; Nowling et al., 2000), Sox9 (Ng et al., 1997), Sox11 (Kuhlbrodt et al., 1998a), Sox18 (Hosking et al., 1995; Ng et al., 1997)). Others displayed no autonomous transcriptional activity in these assays (Sox5 (Denny et al., 1992), Sox10 (Kuhlbrodt et al., 1998b), Sox14 & Sox21 (Uchikawa et al., 1999)) but when natural promotors were used with full length proteins regulatory capacity was observed demonstrating the importance of promotor/enhancer context for appropriate activity *in vivo*.

Modification of the local DNA geometry of enhancer sequences *in vitro* has various effects upon reporter expression but this is only one of the ways in which

these proteins are able to influence gene activity. Mutations introduced into the promotor of the myelin protein zero gene (P_0) increased the bending angle that bound *Sox10* was able to induce (by 20 degrees) but this had little effect on transcriptional activity *in vitro* unless homodimer formation was also interrupted (Schlierf et al., 2002). Cooperation is also key to the optimal activation of the *Fgf4* enhancer by SOX2 and OCT3/4 although mutations that reduced the bending angle by half actually increased reporter activation three fold. Abolishing the bendability of the DNA itself did not interfere with the ability of SOX2 to bind its target but did abolish transactivation activity. The authors argued that the absolute angle of bend is not as important as the direction (phase) of the bend (Scaffidi and Bianchi, 2001). In both these cases a particular geometric arrangement of promotor/enhancer DNA is likely to aid cooperative interaction between proteins that associate with it and in this way facilitate an optimal level of gene regulation.

1.7 Strategies have evolved to ensure selective regulation despite binding similarities.

Binding of Sox family members to the same consensus sequence creates the possibility that any Sox protein might bind to another's target. The primary mechanism that limits the effect of transciptional regulators is to restrict the sites where these genes are expressed. Sox genes are expressed in highly ordered patterns in diverse tissues in many organisms and restricted patterns of expression during development, in particular, indicate that these proteins have important roles in cell fate decisions. However expression of several Sox proteins in the same cells at the same time is often observed (Cheung et al., 2000; Lefebvre et al., 1998; Montero et

al., 2002; Shen and Ingraham, 2002; Uwanogho et al., 1995). Where there is overlapping expression of Sox genes from the same sub-group questions of functional redundancy arise whilst an overlap of more diverse Soxes may suggest some differential regulation of the same targets mediated through common consensus sequences.

To achieve meaningful and effective regulation of target genes, selective mechanisms exist so that any single family member only influences a relevant set of genes and these proteins employ a number of strategies in order to restrict any promiscuous influence upon gene expression. Auto-inhibitory activity, that prevents the close association of protein to target and cooperative binding with other transcription factors play significant roles in curbing out of context regulation (Wiebe et al., 2003). SOX11 is unable to bind the Sox consensus sequence unless an auto-inhibitory acidic region in its C terminus is first neutralized, but this mechanism is perhaps limited to group C genes that also have an acidic domain (Wiebe et al., 2003). In the same assay, SOX2 was able to bind target DNA but needed to cooperatively bind OCT3/4 for maximal activation of the reporter, thus demonstrating a more common mechanism of target specification (Kamachi et al., 2000). Similarly *in vivo*, SOX9 activates *Col2a1,* during chondrogenesis, however ectopic expression of *Sox9* only results in the activation of *Col2a1* in a subset of the extra sites indicating that further factors must be necessary for selective expression of *Sox9* targets (Bell et al., 1997).

In vitro the affinity of Sox proteins for their DNA targets is not especially strong (Kd 10^{-8}-10^{-9}M) when compared with good activators such as GAL4 that binds

to its target two or three magnitudes stronger (Kd 10^{-11}M) (Kamachi et al., 2000). Although *in vitro* binding to consensus sequences of many of the Sox proteins can be achieved this may not represent the *in vivo* situation where specific partnerships with other DNA binding proteins not only play a crucial role in stabilizing their association with target elements, but may also dictate their regulatory influence (Wilson and Koopman, 2002).

There is a strong association with members of homeobox containing gene families such as Pax and POU domain proteins (Wilson and Koopman, 2002). The combination of these two classes of sequence specific transcription factors is perhaps one that has been conserved and refined over a very long period as their ancestral genes also show synergistic DNA binding and activation activity (Dailey and Basilico, 2001). Zappavigna et al. 1996showed that the direct interaction of HMG1 with HOX proteins (sequence specific homeobox genes) was able to increase the amount of HOX-DNA complex formed *in vitro* and HOX mediated transcriptional activation *in vivo (Zappavigna et al., 1996). Fgf4* was one of the first genes shown to be synergistically regulated by the specific interaction of SOX2 and the POU domain protein OCT3/4, which bind to adjacent sites of its enhancer (Yuan et al., 1995) with similar associations subsequently found necessary for the regulation of other genes such as UTF1 (Nishimoto et al., 1999), HOXb1-ARE (Di Rocco et al., 2001), *Fbx15* (Tokuzawa et al., 2003) and *osteopontin* (Botquin et al., 1998).

Specific partnerships not only offer a mechanism for selective regulation of target genes by Sox proteins (reviewed by Kamachi 2000) but also a way in which alternative levels of gene expression may be achieved through the same regulatory

element (Kamachi et al., 2000). *Sox11* and *Brn-2* have overlapping patterns of expression with *Sox2* particularly in the developing CNS (Uwanogho et al., 1995). In an assay similar to that employed to show a specific partnership between SOX2 and OCT3/4 that confirmed their activation potential, SOX11 was demonstrated to be a sixty times more potent activator of the same reporter, either alone or when co-expressed with OCT3/4 (Wiebe et al., 2003). Activation of this reporter by SOX11 was increased further when BRN-2, another POU protein, was introduced instead of OCT3/4 (Wiebe et al., 2003). This along with other examples demonstrates *in vitro* how differential regulation through a common regulatory element might be achieved by the association of alternative partners and how this might recapitulate the situation *in vivo* (Kuhlbrodt et al., 1998a; Kuhlbrodt et al., 1998b).

In some instances partners may be other Sox proteins. Homodimerization of SOX10 (group E) ensures that it can stably associate with a pair of target sequences that an undimerized SOX protein is unable to bind and in this way restricts the activation of the myelin gene protein zero (Peirano et al., 2000). Homo or hetero dimerization also appears to be important for the regulation of the *Col2a1* gene by L-SOX5/SOX6 (group D) (Lefebvre et al., 1998).

An important question, however, is what happens *in vivo* where several possible combinations of partners may occur simultaneously and so context is therefore extremely important. SOX21 represses SOX1 activation of the DC5 crystallin reporter in lens cells and has no intrinsic activation potential, whilst in a neuroblastoma cell line, SOX21 was able to activate a reporter containing a distal

regulatory element of the μopioid receptor (Hwang et al., 2003; Uchikawa et al., 1999).

Different modes of transcriptional regulation may be achieved by the assembly of multi-protein regulatory complexes. Differential regulation of a gene via the same regulatory element reflects the makeup and/or stability of the multi-protein complexes that assemble on it which is perhaps why even very similar (Sox) proteins may elicit different responses in the same context (Kamachi et al., 1998). Cooperative interaction between the Soxes and other DNA binding proteins dictates a specific association with DNA. This perhaps provides a foundation for the recruitment of accessory factors through direct protein interaction or architectural changes to the DNA. This offers a mechanism whereby significantly more complex regulatory networks may have arisen by the alteration or addition of relatively few components and is perhaps why there is an association between the evolution and expansion of families of regulatory genes, such as the Soxes, and the increased complexity and diversity of organisms.

1.8 Gene duplication and the evolution of complex regulatory networks

In 1970 Ohno proposed that diversity and complexity was facilitated by gene duplication and functional redundancy arguing that this enables genes to accumulate modifications that would otherwise be deleterious (Ohno, 1970). In the case of the invertebrate Sox family, one level of complexity seems to have been achieved by the evolution of a small group of transcription factors with diverse function but united by a common DNA binding domain. Vertebrates however have a far more extensive set

of Sox genes often with numerous orthologues to each of the invertebrate Sox genes (Fig. 1.1 & Fig. 1.2).

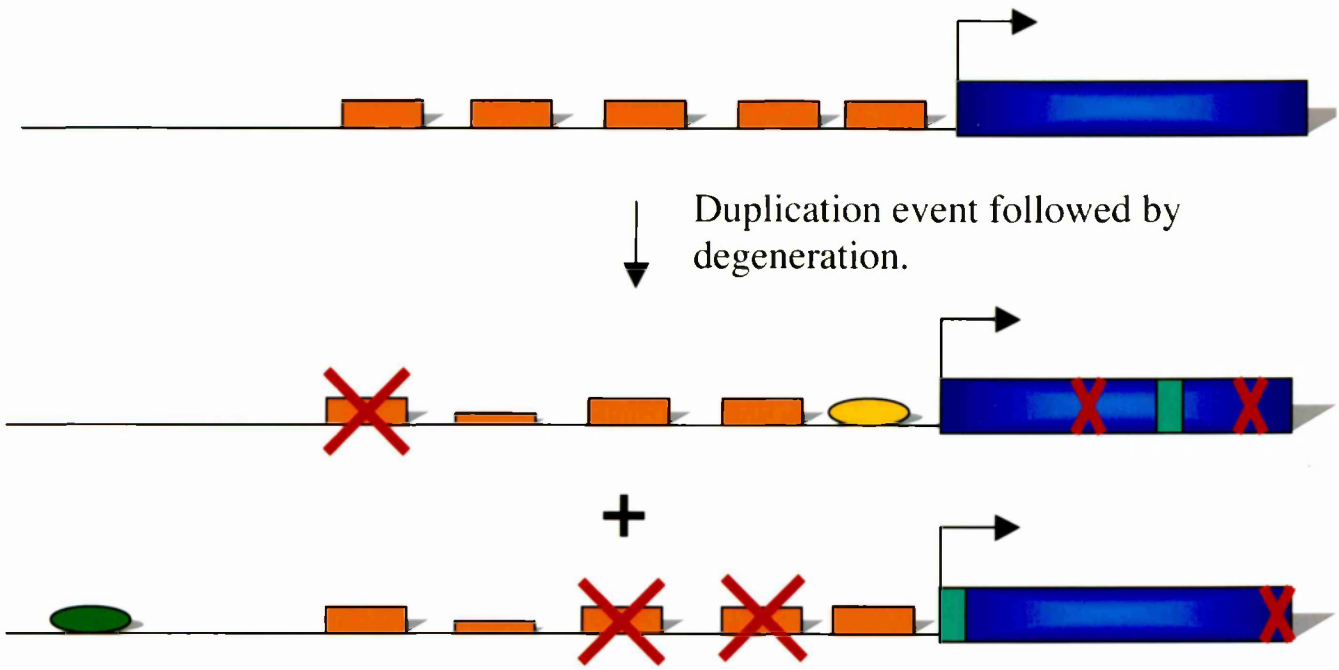
The conservation of similar transcription factors offers the possibility of a wide range of responses by any given cell in which they are expressed. A single transcription factor is likely to regulate a cohort of targets through its association with a number of other factors. Whilst its closest relatives may be capable of regulating a very similar group of targets the slight differences between them are likely to produce altered affinities with certain cofactors which in turn may alter various aspects of target gene regulation. In combination, these factors have the potential to produce a wide range of regulatory outcomes but the fine tuning of a cells regulatory network, and thus its ideal identity, may be key to their evolutionary conservation.

The classical model for the evolution of duplicated genes describes how once duplicated one of the pair will rapidly accumulate deleterious mutations (within a few million years) and will only be conserved if it acquires a novel function. But this does not explain why so many duplicated genes seem to have been conserved in higher organisms. The duplication-degeneration-complementation (DDC) model, however, describes how gene duplicates may be conserved if regulatory mutations render each copy essential so overall their expression is (at least) equal to that of the ancestral gene (Fig. 1.3) (Force et al., 1999). In time this could either result in functionally identical genes being expressed in mutually exclusive sites or very similar, but perhaps not identical, genes being expressed in the same cells at the same time. In the latter case these genes, which in a given context, encode essentially the

same proteins may achieve an optimal dose through regulatory mutation but subtle functional differences could also account for simultaneous expression.

Figure 1.3 Duplication and evolution of gene families

The DDC model after Force 1999. When a gene is duplicated it is released from the normal pressures that conserve its sequence. Gene pairs accumulate mutations in regulatory (orange boxes) and functional sequence (blue boxes) as long as the overall role of the ancestral gene is maintained. Mutations imparting novel function (green boxes) or expression sites (yellow lozenges) may also be accumulated resulting in a pair of closely related but unique proteins that may have some overlapping function. (Regulatory elements may possibly be assimilated during the duplication event).


Duplication event followed by
degeneration.

Within the Sox family there are examples of different strategies of specialisation acquired following duplication events this includes diversification of function of similar proteins expressed in the same cells. L-SOX5 and SOX6 are in sub-group D and are 67% identical to each other and 94%identical in the HMG domain (Kamachi et al., 2000; Smits and Lefebvre, 2003). Both are expressed in all cartilages and they can dimerize alone or in combination. Because of their similarity, it was hypothesised that there might be functional redundancy between these two genes. *In vivo,* a severe disruption in chondrogenesis was seen only when mice were null for both of these genes, being lethal at 16.5 days post coitum (dpc). The individual knockout animals survived past birth but each had slightly different phenotypes that caused lethality before they were weaned, whereas compound heterozygotes were viable and fertile (Smits and Lefebvre, 2003). This suggests that although some aspects of chondrogenesis can be performed by the remaining protein not all functions can be fulfilled, even by a very similar counterpart.

A different resolution appears to have been attained in the case of the group E genes *Sox9* and *Sox10* that are very similar to each other (72% amino acid similarity (Schepers et al., 2000)) but are expressed in mutually exclusive patterns. Null mutation of either of these genes in mice results in severe developmental phenotypes. However, another group E gene *Sox8* is expressed in a pattern that is essentially the sum of that of *Sox9* and *Sox10,* but it does not seem to have a similar level of functional responsibility as only relatively mild phenotypes are observed when it is deleted (Sock et al., 2001). *Sox8* may demonstrate the importance of more subtle contributions as deletion does lead to minor phenotypes that affect an individuals ability to thrive. Whilst some phenotypes could be attributed to absence from cells

where there is no overlapping expression of other group E genes there are abnormalities in other structures, such as the tarsals of the hind limbs, where there is overlapping expression of *Sox9*. These data suggest that whilst some functional conservation is exhibited between subgroup members certain activities clearly provide a unique contribution to an individual.

Where there is overlapping expression of a gene with that of very similar family members, redundancy is often heralded as the reason for the apparent lack of obvious phenotypes from the null mutation. However, there are at least two alternative explanations. First, the phenotypes may only be revealed by minute analysis of specific regions and molecular markers. Secondly phenotypes may not be manifest under normal laboratory conditions. This can be demonstrated in yeast whose environment can be readily manipulated. In this system an examination of the effect of null mutation reveals that increasing the number of growing conditions decreases the chance that any given null mutation will have little or no effect (Gu et al., 2003). In higher animals, such as mice, this effect is more difficult to ascertain and it may be necessary to devise specific assays in order to reveal the function of certain genes.

1.9 Sox *group B* genes:similarities & differences.

SoxB genes were the first members of the *Sox* family to be discovered and represent the subgroup that has been most extensively studied across a range of animals (Gubbay et al., 1990). As well as sharing sequence similarity, all group B genes so far isolated are expressed in the developing nervous system although individual members can also be detected in a spectrum of other tissues during development. Similar sequence and expression characteristics observed between

distantly related species indicates that certain regulatory roles and relationships may have also been conserved and highlight the importance of functional analysis in lower organisms.

1.10 Vertebrate versus Invertebrate Sox genes functional similarities.

In the fruit fly, Drosophila, the *SoxB* gene *Dichaete* (SOX70D, *fish, fish-hook* or *SoxB2.1*) is required for the correct formation of the midline structures and phenotypes arising from its disruption can be rescued by expression of mouse *Sox2* (Soriano and Russell, 1998). *Dichaete* shares 88% identity in its HMG domain to SOX2 as well as limited identity within the C terminal domain but may actually represent an ancestral form of the entire sub-group B. Removal of C terminal sequence produces a truncated form of *Dichaete* that is able to rescue midline phenotypes as efficiently as the full length protein but lacks activity when ectopically expressed in other sites of *Dichaete* expression, such as eye-antennal or wing imaginal disks (Mukherjee et al., 2000; Soriano and Russell, 1998). This suggests that *Dichaete* has discrete functional characteristics with context dependent importance. Midline structures rely most heavily upon some property of the *Sox* HMG domain but the C terminal domain possesses significant but uncharacterised functional properties highlighted during gain of function experiments.

Conserved interaction with POU partners

Correct formation of midline structures, especially midline glia, is also dependent upon the POU domain protein *Ventral veinless* and genetic experiments implicate interaction with *Dichaete* in the regulation of target genes (Soriano and Russell, 1998). These experiments further support an ancient relationship between

SOX and POU proteins that have been shown to cooperate via interaction between their DNA binding domains (Kamachi et al., 2000). Sox genes do display partner preference but can cooperate with other proteins of the same class that, at least *in vitro,* can result in different activation levels of the same reporter constructs (Wiebe et al., 2003).

A second Drosophila gene *SoxNeuro* (*SoxN, SoxB2.1*) is even more similar to mouse *Sox2* than *Dichaete* with 92% amino acid identity in the HMG domain although again there is limited similarity in flanking regions (Sasai, 2001). Like *Dichaete, SoxN* is also expressed in the developing nervous system with a degree of overlap between these two related proteins and *SoxN* mutants have the strongest phenotypes in lateral CNS where its expression does not overlap with *Dichaete* (Cremazy et al., 2001; Overton et al., 2002). In areas where there is overlapping expression, neural phenotypes are fairly mild in both *SoxN* and *Dichaete* mutants. Not surprisingly simultaneous disruption of both Drosophila *SoxB* genes results in general neural hypoplasia and indicates a degree of functional equivalence in some contexts. However in other situations these proteins have different activities and even antagonistic effects. This apparent contradiction may be explained by differences in the nature of the cooperative interactions required for target selection and regulatory activity. Where this depends upon partner interaction via the HMG domain, *Dichaete*, *SoxN* and even mouse Sox2 may functionally compensate for each other. Targets where there is a requirement for interactions not involving the HMG domain are likely to recruit more specific combinations of factors that will in turn elicit different regulatory activities. Cell identity may even depend upon competition for common targets and may explain some of the peculiar responses to ectopically

expressed protein (Soriano and Russell, 1998). The evolution of higher organisms may have exploited some of these regulatory principals with the creation of larger groups of similar proteins providing the opportunity of further, subtle, modification of the regulatory network.

1.11 Vertebrate group B sequence similarities and differences.

Within the HMG box vertebrate group B proteins, from the mouse and chicken, are greater than 94% similar to each other, but a further subdivision separates group B1 (*Sox1 Sox2* and *Sox3)* from group B2 (*Sox14* and *Sox21)* as there is only limited similarity between them in non HMG domains (Kamachi et al., 1998; Uchikawa et al., 1999) (Fig. 1.2).

When non-HMG domain sequence from the three mouse B1 genes is compared it is clear that despite significant similarity each has a unique combination of structural motifs (Collignon et al., 1996). A PRD-type repeat (His-Pro) is a unique feature present only in SOX1, whilst polyalanine repeats, two of which are in similar positions, punctuate regions of homology in SOX1 and SOX3. Comparison with chicken sequence reveals conservation of domain arrangements except that poly-alanine repeats are absent in chicken SOX3. When homologues from chicken and mouse are compared then identity is much higher with SOX1, SOX2 and SOX3 sharing 87%, 93% and 70% amino acid identity respectively despite the closest common relative occurring more than 310 million years ago (Kamachi et al., 1998). Chicken SOX14 and 21 also show a high degree of non-HMG box homology (about 50%) and again regions of homology are separated by stretches of poly alanine repeats (Uchikawa et al., 1999).

Such extensive similarities would suggest functional conservation between the most closely related genes. Indeed SOX1, SOX2 and SOX3 have all been implicated in the in the consolidation of neural fate and are able to activate a luciferase reporter via the δ1-crystallin enhancer in lens cells (Graham et al., 2003; Kamachi et al., 1999; Kamachi et al., 1998; Uchikawa et al., 1999). However there is almost a three fold difference in the level of activation of this reporter construct by these different proteins indicating that the differences that do exist significantly effect their activity (Kamachi et al., 1999; Kamachi et al., 1998).

Two *group B2* genes *Sox14* and *Sox21* have been shown, in similar assays, to act as repressors, however, their levels of repressive activity differ (Uchikawa et al., 1999). Repressive activity is conferred by their carboxyl terminal domains and can also specifically interfere, *in vitro,* with SOX1 activation of the DC5 region of the crystallin enhancer in lens cells (Uchikawa et al., 1999). *In vivo, Sox14* expression phenocopies the effect of a dominant negative form of *Sox2* in the chick neural tube further supporting a repressive role (Graham et al., 2003). It would be misleading to label the *Sox B2* genes exclusively as repressors since *Sox21* was subsequently isolated in a one hybrid screen using the distal promotor of the μopiod receptor as bait and in this context shown to be a transcriptional activator (Hwang et al., 2003).

1.12 Expression of Sox group B1 genes

As well as displaying significant sequence similarity *Sox1, Sox2* and *Sox3* display considerable overlap in their expression patterns. These genes are expressed predominantly in the developing nervous system. However de-novo RNA expression has also been detected in other tissues in the developing embryo and adult, although

most analysis has only been carried out on embryos up to mid gestation. Expression is detected in a variety of cells with different developmental potential, but in general it is suggestive of a role in pluripotent or undifferentiated cell types. *Sox2* RNA is the first to be detected, at about 3.5dpc, in the inner cell mass of preimplantation embryos (Avilion et al., 2003). At the pre-streak stage (6.5dpc) *Sox2* and *Sox3* are expressed throughout the epiblast and extraembryonic ectoderm then in the initial stages of gastrulation expression becomes restricted to anterior ectoderm and the chorion. *Sox3* expression is rapidly down regulated in the extraembryonic component whilst *Sox2* continues to be expressed at the headfold stage (Wood and Episkopou, 1999). Beginning at the late streak stage, *Sox3* ectoderm expression spreads toward the posterior so that by the headfold stage *Sox2* and *Sox3* have formed complementary anterior posterior gradients. It is at this stage when a low level of *Sox1* expression can first be detected in the neural plate ectoderm. All three genes are expressed in the neuroectoderm from the onset of somitogenesis and show varying degrees of overlapping expression (Collignon, 1992; Wood and Episkopou, 1999). *Sox2* appears to have an expression pattern which almost entirely encompasses that of the sum *Sox1* and *Sox3* except for a few areas of the hindbrain where some complementary expression patterns are seen. There are a few regions, however, such as the floor plate and future roof plate of the neural tube that have unique expression of *Sox2*. Cells of the peripheral nervous system display expression of *Sox2* but no expression has been observed in neural crest cells that give rise to it (Wood and Episkopou, 1999). In the spinal cord at 13.5dpc, overlapping expression of *Sox1* and *Sox2* continues to be confined to the ventricular zone, with *Sox2* alone being expressed in

the floor and roof plate regions *Sox3* expression is also largely confined to the ventricular zone (Rizzotti, unpublished data).

Until midgestation *Sox1* and *Sox2* expression is restricted to neuroectoderm but from about the six somite stage, *Sox2* is also observed in a broad region of surface ectoderm from which all sensory placodes will derive. Within the latter, it becomes restricted to the individual placodes as they appear (Collignon, 1992). Overlapping expression of *Sox3* can be seen in ectoderm adjacent to the hindbrain and overlying the second branchial arch at these early stages. Slightly later both *Sox2* and *Sox3* can be detected in the nasal placodes but only *Sox2* alone is seen in the otic placode (Wood and Episkopou, 1999). *Sox2* is expressed in the earliest stages of lens formation but then gives way to *Sox1* as the predominant gene expressed whilst *Sox3* is not observed in this structure (Kamachi et al., 1998). *Sox2* displays de-novo expression throughout the endoderm of the gut and lung with a partially overlapping expression pattern with *Sox3* that is detected in the posterior part of the foregut (Collignon, 1992; Kamachi et al., 1998; Wood and Episkopou, 1999).

In chick, expression of the three *Sox group B1* genes is largely similar to that of the mouse although there a few notable differences where opposite patterns appear to have been adopted by *Sox2* and *Sox3* such as the epiblast and germ cell lineage (Uchikawa et al., 1999). This might indicate a random evolutionary choice between two very similar genes or perhaps reflect subtle functional differences with different suitability in the respective species. Comparing *SoxB* expression in the developing chick CNS reveals a complicated pattern of expression where different levels and combinations of these genes mark discrete regions (Uchikawa et al., 1999). It is

therefore possible that proper development of particular populations of cells depends on the correct ratio of these factors despite their apparent similarity to each other.

1.13 Disruption of the *SoxB1* genes.

In mice, null mutations in all the members of the *Sox group B1 genes* (*Sox1* *Sox2* and *Sox3*), have been produced with all three models displaying distinct phenotypes upon complete removal of the respective gene. Although there is extensive overlap in the expression pattern of these three genes the most severe phenotypes are observed in tissues where only a single *SoxB1* gene is normally detected. This suggests that in the sites where there is overlapping expression there is some degree of functional compensation with normal expression of other closely related proteins sufficient to support the basic integrity of the regulatory network. It is likely, however, that although these closely related genes share some common function, additional roles have evolved that have assured their continued existence and co-expression.

1.13.1 Mouse knockouts to date – Sox1

Sox1 does not appear to be essential for embryonic viability in mice but the targeted deletion of *Sox1* (*Sox1*$^{m1-/-}$) causes malformations in both eye and brain development (Malas et al., 2003; Nishiguchi et al., 1998). Micropthalmia is observed in mutant animals and is a result of the failure of lens fibre cells to elongate properly, probably as a result of the almost complete absence of γ-crystallins. CNS development appears grossly normal despite widespread expression throughout the neuroepithelium but a major developmental defect in the ventral telencephalon is observed in SOX1 deficient animals. Malformation of the ventral striatum and the

olfactory cortex appear responsible for spontaneous seizures in animals over six weeks and result in secondary malformations of the hippocampus and the CA1 pyramidal cell layer and premature death (Malas et al., 2003).

1.13.2 Mouse knockouts to date – Sox3

Targeting of the X-linked *Sox3* gene in XY ES cells produced clones that were completely null as the only copy of the gene present was mutated (Parsons, 1997). The production of chimeras from these cells results in embryonic lethality when the ES cell contribution is over about 10%. Embryos recovered at 9.5dpc showed severe aberrant morphology with posterior, mesodermally derived structures most severely affected. *Sox3* is not expressed in these tissues and it is thought that these phenotypes are a consequence of an abnormality in cells as they pass through the primitive streak. More recently a conditionally null mutation of *Sox3* has been produced which does not show the same phenotypes. Instead, mice completely null for *Sox3* can be viable but often display a range of abnormalities associated with abnormal formation and function of the pituitary gland (Rizzoti et al., 2004).

1.13.3 Mouse knockouts to date – Sox2

The null mutation of *Sox2* was created in ES cells by replacing the coding region of this gene with the selection/reporter cassette *ßgeo* (Avilion et al., 2003; Friedrich and Soriano, 1991). Heterozygous ($Sox2^{\beta geo-/+}$) mice produced from targeted clones are usually viable and show little phenotype, whilst the homozygous null mutation ($Sox2^{\beta geo-/-}$) is always lethal at the peri-implantation stage at around 5.5dpc. Survival to this stage may be the consequence of a pool of stable, maternally derived, SOX2 which appears to persist in pre-implantation embryos and suggests that one

contribution this protein makes is to help maintain the pluripotency of embryonic progenitor cells. Chimera rescue experiments revealed a cell-autonomous requirement for zygotic SOX2 around 5.5dpc by cells of the epiblast, but rescued embryos then died shortly after 7.5dpc due to an extraembryonic defect (Avilion et al., 2003). A similar failure of inner cell mass/epiblast lineages is observed in embryos null for the POU protein OCT3/4 supporting evidence of cooperative interaction between these two proteins in the regulation of a number of genes expressed at these stages including *Fgf4* (Ambrosetti et al., 1997), *osteopontin* (Botquin et al., 1998), UTF1 (Nishimoto et al., 1999) and *Fbx15* (Tokuzawa et al., 2003) and maybe *Sox2* itself (Tomioka et al., 2002). Failure at these stages of development makes phenotypic analysis technically demanding and it is impossible to obtain information about the role of SOX2 at other sites later in development.

These data demonstrate the significant role for *Sox2* in a number of tissues prior to 7.5dpc. Isolation of purified neural precursors on the basis of *Sox2* expression (Li et al., 1998), as well as the detection of further *Sox2* expression sites throughout mouse development suggest a significant role for *Sox2* not only in these stages but throughout embryogenesis, but early lethality precludes the use of $Sox2^{\beta geo-/-}$ embryos to investigate contributions later in development.

1.13.4 Summary of the *SoxB1* null mutations in mice

An examination of the mice null for *SoxB* genes indicates that aberrant phenotypes are most commonly manifest in tissues that do not have overlapping expression of *Sox1*, *Sox2* or *Sox3*. Lens malformation, observed in $Sox1^{m1-/-}$ mice, seems to coincide with a stage in normal lens development when *Sox1* is up-regulated

and *Sox2* is down-regulated (Nishiguchi et al., 1998). The mutant embryos produced with *Sox3* null ES cells are indicative of an abnormality in ectodermal cells that are destined to form mesoderm but no longer express *Sox2* and have never expressed *Sox1*. *Sox2* is the first of these genes to be expressed in the embryo and so mice null for this gene show the earliest abnormal phenotype of all, resulting in lethality. Rescue of this phenotype reveals further abnormalities resulting from a different site of unique *Sox2* expression (Avilion et al., 2003). Together these data suggest all three genes have critical functions in development, but that there is likely to be some degree of overlapping function.

Expression analysis and functional studies in other organisms suggest that the *SoxB1* genes have important roles in the early stages of neuronal development. In order to reveal these roles in mice, null mutations have been created, but these models throw little light on gene function due to early lethality or possible compensation by the overlapping expression of closely related genes. The complete disruption of *Sox2* or *Sox3* specifically in the developing nervous system and avoiding early lethality is now technically possible but may still expose little of their function due to compensation by the remaining wild-type *SoxB1* alleles. If these genes do have equivalent function the role of the *SoxB1* genes will only be revealed by the overall reduction in *SoxB1* dose. The disruption of multiple alleles may therefore be able to reveal some of the functions these genes have to play in neuronal development.

1.13.5 Human mutations involving *SOXB1* genes: *SOX2.*

Evidence regarding the role of the *SoxB* genes can also be elucidated by studying the effects of genetic lesions in humans. Fantes et al. 2003 show four

patients with anophthalmia (two bilateral and two unilateral anophthalmia with contralateral micropthalmia) that are heterozygous for *de novo* truncation mutations in *Sox2*. Interestingly the more severe of the patients have non sense mutations in the C-terminal of *SOX2* whereas the slightly less severe mutations in the middle of the DNA binding domain and should disrupt all binding. It may be possible that the more severe mutations are also acting in a dominant negative fashion that interferes with the function of other *SOX* proteins.

1.13.6 Human mutations involving *SOXB1* genes: *SOX3*.

A number of human patients have been characterised as having some degree of X linked mental retardation (XLMR) in connection with the Xq26-Xq27 region where (the X linked gene) *SOX3* is located (Laumonnier et al., 2002). Patients carrying a chromosomal inversion, including *SOX3* are mentally retarded but display no significant physical abnormalities. Short stature and mental retardation in N3 patients is linked to a *SOX3* with eleven additional alanines in one polyalanine tract (711-783 dup (33bp)) roughly doubling its length. Two patients (from the same family) carrying the T84 mutation, where nine alanines are deleted from the same tract (718 del 27), were severely mentally retarded. Their maternal grandfather however carried the same mutation but showed no signs of similar mental retardation suggesting an additional mutation carried by the boys. Alanine expansions have been found in a number of other genes linked with human disease with a suggestion that they may cause a suppression of transcriptional activity (Han and Manley, 1993). T19 were short, microcephalic with mild mental retardation and peculiar facies but

had no obvious abnormality in *SOX3*, although XLMR was mapped to the Xq26-Xq27 region (Laumonnier et al., 2002).

1.14 Scope of thesis

When one is confronted by a large group of similar regulatory genes that display overlapping expression patterns it is important to elucidate the role that each one plays. Three reasons why coexpression may be found are the following. Firstly redundancy. In this instance there is expression surplus to requirement and removal of one gene will have no effect in regions of overlapping expression. Secondly all members of the group are functionally equivalent but an optimal level of activity is achieved by the cumulative contribution of the expressed genes. Finally the slight differences in gene sequence between each of the members dictate distinct regulatory activities that could allow differential regulation of the same or discrete targets.

A suitable strategy to distinguish between each of these possibilities would comprise several experimental steps. In this instance I have analysed the expression of *Sox2* in the mouse embryo as well as carrying out comparative analysis with *Sox1* and *Sox3* in specific regions. This will highlight the areas most likely to be affected by the null mutation of any one of these genes. Phenotypes arising from the null mutation of individual *SoxB1* genes tend to be exhibited in sites where there is unique expression and in the case of the *Sox2* null precludes any investigation after 5.5dpc. In order to study *Sox2* by null mutation at later stages a strategy avoiding early lethality must be adopted. I attempted to achieve this by both the creation of a targeted mutation allowing a temporally and spatially specific ablation of *Sox2*, as well as identifying randomly generated *Sox2* regulatory mutants. In order to address

the role of *SoxB1* proteins in regions of overlapping expression the null mutation of individual genes is a difficult starting point, therefore I created mice carrying compound null mutations using the existing animal models. By the creation of animals harbouring all the various combinations of available *SoxB1* null alleles an idea of their relative contributions may be ascertained. The introduction of multiple null alleles will reduce the overall *SoxB1* dose to a level that might reveal more obvious phenotypes and illuminate the contribution these genes make during development. From this an assay may be developed sensitive enough to examine the role the individual genes might have in these sites of overlapping expression. A reduction in the overall dose of these genes does not, however, allow that one to distinguish between the second hypothesis of functional equivalence and the third hypothesis of distinct regulatory activity. In order to achieve this *in vivo* a series of targeted replacement experiments would be necessary to see if each gene under the control of another's regulation can perform in exactly the same way.

Chapter 2 Expression of *Sox2* during mouse embryogenesis.

2.2 Introduction.

There is reiterative use of specific critical signalling molecules throughout development. Notable examples would be fetal growth factor 4 (*Fgf4*) and sonic hedgehog (*Shh*) which make significant contributions during CNS development and morphogenesis of the limb. The influence that such molecules have on a given cell is mediated via further proteins such as membrane receptors and transcription factors, the exact combination of which will dictate a particular cellular response. Further analysis might reveal diverse regulatory strategies that are built around a foundation of common components. It is unlikely that a single transcription factor might elicit an identical response from all its targets at different spatial and/or temporal points. Rather, a contribution is made to an overall response that is unique to a given cell type or even individual cell. This is achieved by the interaction of regulatory proteins producing a network flexible enough to provide the subtlety and diversity of gene expression that organisms require, especially those with greater complexity. Although similar mechanisms may be employed in the regulatory networks of diverse cell types, expression and functional data indicate that, it is the particular combination of all these components that defines cellular identity. It may also be true that similar regulatory proteins, like those within highly related families, might offer a mechanism whereby gene expression levels and thus cell behaviour can be modified in a subtle way without drastically disrupting the core regulatory network. Expression data is thus vital in order to understand some of the complexities of gene regulation within a particular cell type and will consequently help to identify common

or unique interactions providing the individual response that defines the identity of any given cell.

The expression pattern of group B *Sox* genes has been widely investigated in the early stages of both mouse and chick development but only a few specific sites of expression have been described at later stages (Avilion et al., 2003; Cheung et al., 2000; Collignon, 1992; Ishii et al., 1998; Pevny and Rao, 2003; Uchikawa et al., 1999; Wood and Episkopou, 1999). A comprehensive description of the expression pattern of these genes, including *Sox2,* throughout development is thus important to fully understand how these proteins function and in order to decipher the phenotypes that arise from manipulating *Sox* expression *in vivo* or *in vitro.* Further analysis of the *Sox2* expression pattern after mid-gestation was carried out in order to extend the current description and to establish whether there are any uncharacterised sites of expression.

2.2.1 *Sox2*$^{\beta geo-/+}$ can be used as a tool to study *Sox2* expression.

When the *Sox2* allele was originally targeted with a null mutation, mice were produced carrying the reporter/selection fusion protein *ßgeo* that replaced 3.5 kilobases (kb) of one of the *Sox2* alleles including the entire coding region (Avilion et al., Fig. 3.2a). *ßgeo* is a fusion between the bacterial gene beta galactosidase (*ßgal*) and the antibiotic resistance gene neomycin (*neo*) (Friedrich and Soriano, 1991). This strategy provided an efficient way in which to isolate correctly targeted embryonic stem cells but had the added benefit that animals made from these cells should express the *ßgeo* protein under the control of the endogenous *Sox2* promoter. Although some regulatory function could potentially have been lost expression from

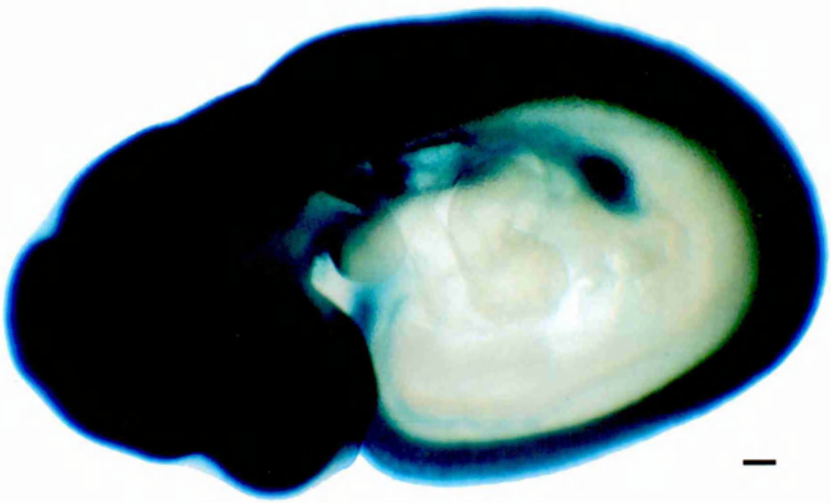
Figure 2.1 *ßgeo* expression in a *Sox2*$^{ßgeo-/+}$ embryo at 9.5dpc.

X-gal staining of Sox2$^{ßgeo-/+}$ *embryos compares favourably with the published expression pattern of Sox2 at 9.5dpc (see Avilion et. al. 2003 figure 1F). Sectioning of stained samples reveals the detail of the expression but reconstruction and three dimensional modelling gives a much clearer impression of staining patterns.*

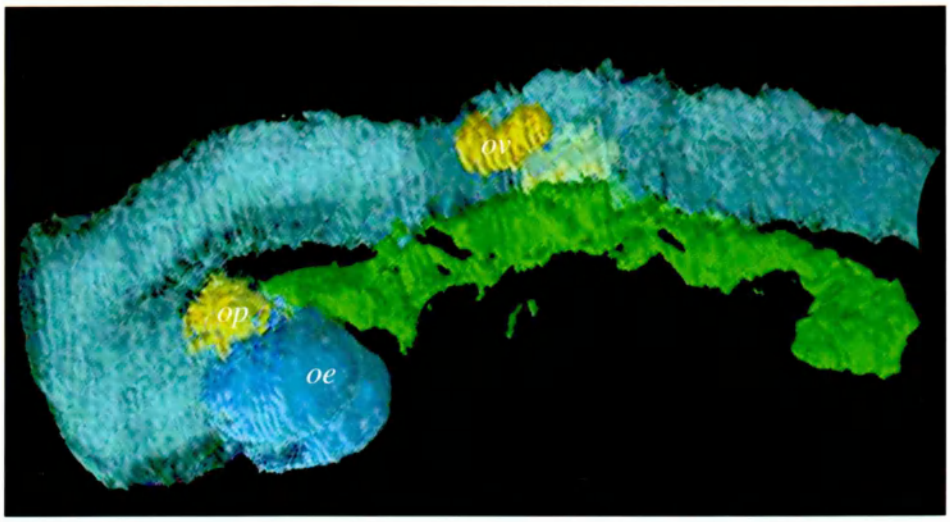
(a). *ßgeo expression is revealed by X-gal staining of 9.5dpc Sox2*$^{ßgeo-/+}$ *embryos. The ßgeo selection cassette, under the control of endogenous Sox2 regulatory elements, produces staining throughout the neural tube as well as in the gut endoderm and sensory placodes that are difficult to visualise in wholemount. Bar is 100µm.*

(b). *Three-dimensional reconstruction of the X-gal staining pattern of a 9.5dpc embryo with pseudo-colouring highlights different components of the Sox2 expression pattern. This model shows only the parts of the embryo that are expressing ßgeo (and are also positive for SOX2 expression) at 9.5dpc. The most extensive staining is seen in the neural tube (transparent cyan) and in the oropharyangeal and gut endoderm (solid green). Staining is seen in all the sensory placodes at this stage. Olfactory epithelium (oe) is solid cyan and surrounds the anterior end of the neural tube. Otic placode (op) lies adjacent to the olfactory epithelium and is coloured solid yellow. Staining in the otic vesicle (ov) is also coloured solid yellow but lies in a more caudal position.*

(a)



(b)



this cassette should reliably mimic the expression pattern of *Sox2* and so can be used to indicate sites of *Sox2* expression that warrant further investigation (Avilion et al., 2003; Tomioka et al., 2002).

2.3 Results

The expression of the *βgeo* fusion protein was visualised using the chromagenic substrate 5-Bromo-4-chloro-3-indolyl beta-D-galactopyranoside (X-gal) in partially fixed $Sox2^{\beta geo-/+}$ tissue (materials and methods). Visualisation of the coloured product of X-gal is a robust and sensitive method that can be used to rapidly assess expression patterns in embryos that are otherwise not accessible to analysis by wholemount mRNA *in situ* hybridisation or immunohistochemistry. Whole mount detection of *βgal* activity was carried out in staged $Sox2^{\beta geo-/+}$ embryos between 9.5dpc and 17.5dpc (Fig. 2.1 & 2.2a-h). These samples were then serial sectioned to reveal the detail of this staining.

Although the detection of *βgal* activity is a useful tool for preliminary analysis, the stability of this bacterial enzyme means that the X-gal staining pattern produced may not be exactly representative of either the transcript or protein of interest. In order to control for the possibility that X-gal staining might be misrepresenting both the spatial and temporal expression pattern of SOX2 a number of sites showing *βgal* activity were examined by section immunohistochemistry using a SOX2 specific polyclonal antibody.

At 9.5dpc, X-gal stained embryos have an expression pattern closely resembling the mRNA *in situ hybridisation* pattern of *Sox2* as previously reported (Avilion et al., 2003; Collignon, 1992; Wood and Episkopou, 1999). This assures confidence that X-gal staining does indeed accurately represent the actual expression pattern of SOX2 in developing $Sox2^{\beta geo-/+}$ embryos. At this stage, the X-gal staining pattern reveals a number of developing organs that have some degree of *Sox2*

Figure 2.1 *ßgeo* expression in *Sox2*$^{ßgeo-/+}$ embryos between 11.5dpc and 13.5dpc.

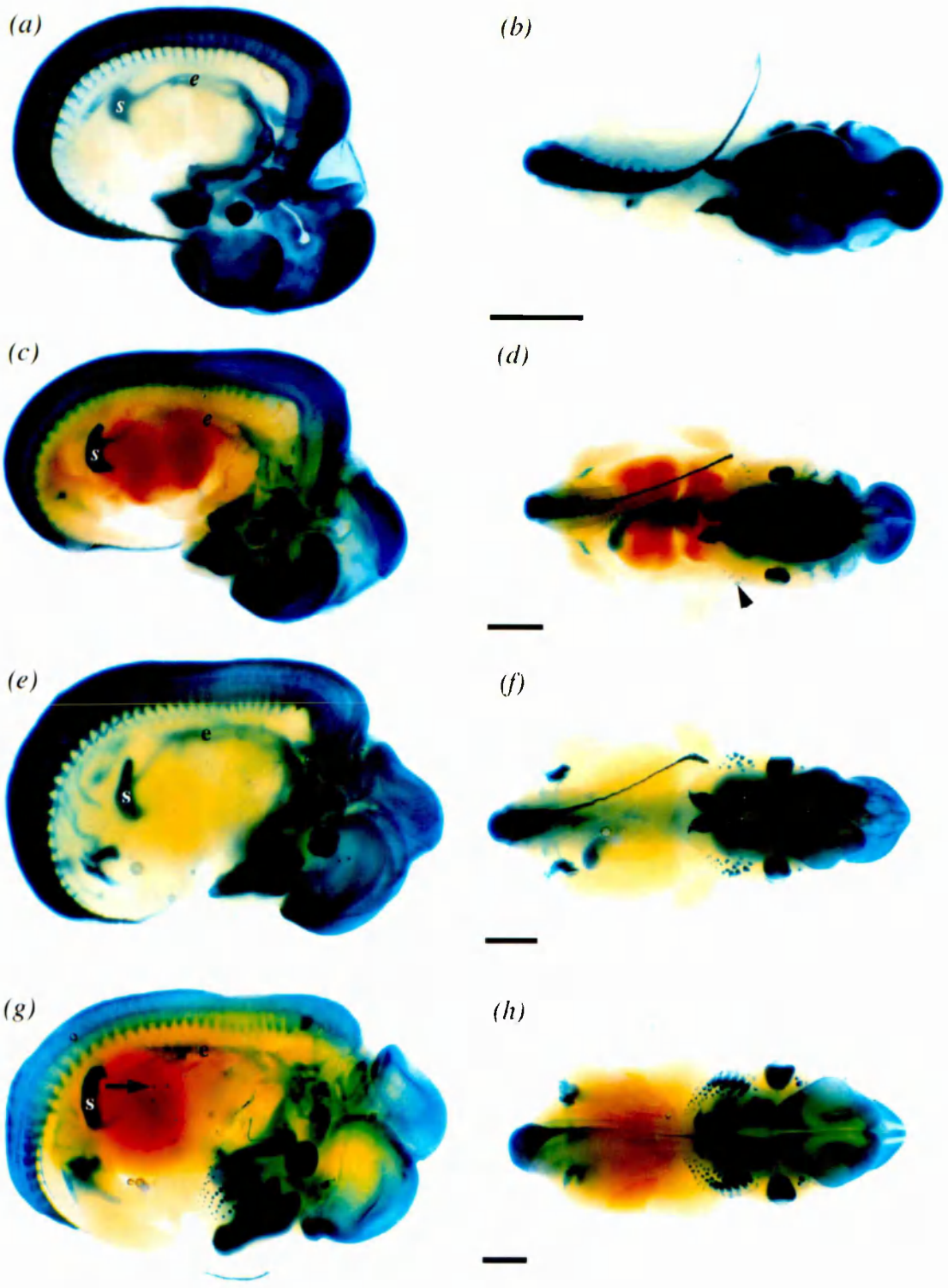
Sox2$^{ßgeo-/+}$ *embryos were harvested at various stages between 11.5dpc and 15.5dpc and subjected to whole mount ßgal staining (materials and methods). Lateral (a,c,e,g) and ventral (b,d,f,h) views of stained embryos are shown. Bar is 1mm in adjacent pairs of pictures.*

(a), (b). *11.5dpc: The majority of staining is in the developing CNS. Endodermal staining of the oesophagus and branching lung shows the initial development of these organs (e).*

(c), (d). *12dpc: At this stage the dermal papilla of the vibrissae are just starting to condense from the surrounding mesenchyme and can be clearly seen lateral to the strong staining in the nasal epithelium (arrowhead).*

(e), (f). *12.5dpc: A few hours later, as the previously induced vibrissal follicles develop, staining of cells in the barrel region of the developing follicles as well as further dermal papillae can be seen. Ganglia of the PNS that will innervate the hind limbs are prominently displayed lateral to the posterior spinal cord. The crescent shaped stomach epithelium marks the posterior limit of gut staining (s).*

(g), (h). *13.5dpc: Staining of the CNS is much weaker as cells differentiate. Most of the vibrissae present in an adult mouse have been induced with lines of whisker follicles marked clearly by blue staining. Follicles of the whisker like hairs above the eye have started to form and are also stained. Spots of staining in the abdominal skin mark condensations of mesodermal cells that are the dermal papillae of the first wave of follicle induction that will eventually form the monotrick hairs in the adult coat (arrow).*


(a)
e
s
(b)
(c)
e
s
(d)
(e)
e
s
(f)
(g)
e
s
(h)

expression that persists throughout embryonic development. The two most significant components of the *Sox2*$^{\beta geo-/+}$ expression pattern can be seen in Fig. 2.1) with extensive staining of the developing central nervous system (CNS) and the endoderm of the emergent gut and lung.

2.3.2 *Sox2* expression in the developing nervous system

Sox2 is expressed throughout the nervous system from the earliest stages of development (Avilion et al., 2003) and persists in the majority of CNS cells until mid gestation (Avilion et al., 2003; Collignon, 1992).

During the second half of embryonic development there is an elaboration on the basic CNS plan resulting from proliferation and differentiation. At 10.5dpc, in whole-mount preparations, X-gal expression is observed throughout the CNS and in the ganglia of the peripheral nervous system (PNS) (Fig. 2.2a,b). *ßgal* expression becomes increasingly restricted from this stage onwards commencing coincidently with the differentiation of the embryonic CNS (Collignon, 1992). Expression is lost first from the ventral neural tube throughout the region of the post-mitotic (motor) neurons. By 13.5dpc a signal is still present but it becomes more refined as further neurons of both the central and peripheral nervous systems differentiate and cease to divide. However, there continues to be populations of cells that maintain *Sox2* expression in the brain and spinal cord up to and after birth (Fig. 2.2g,h).

As the complexity of the developing CNS increases it is necessary to identify the sites of SOX2 localisation on a cellular level. Hence immunohistochemistry was performed on serial sections of mid to late gestation brain, using the SOX2 antibody,

which emulates the pattern of mRNA *in situ hybridisation* results (Collignon, 1992) and *ßgal* staining in the early stages of CNS development,

At 10.5dpc the SOX2 antibody marks the majority of neuroepithelial cells of the developing CNS where neuro and gliogenesis is taking place (Fig. 2.4a). At later stages staining remains associated with populations of undifferentiated and dividing cells that line the ventricles but further populations are also marked by SOX2 expression. At 16.5dpc SOX2 is predominantly visualised in the ventricular and sub-ventricular layers. In the forebrain, cells travelling to the olfactory bulb from the lateral ventricles form a structure called the rostral migratory stream (RMS) that continues to contain neuronal precursors into adulthood (Hartfuss et al., 2001). SOX2 is detected in cells of the RMS but also in those migrating from the ventricular zones of the lateral ganglionic eminence to populate the nucleus accumbens and the olfactory tubercle. The most closely related *Sox* proteins SOX1 and SOX3 were also detected in serial sections showing that in this region there is extensive overlap in their expression patterns. In the forebrain, staining in the ventricular zones and RMS always appeared very similar, but SOX1 staining in the olfactory tubercle appeared to be slightly more extensive than either SOX2 or SOX3 (Fig. 2.3a-c) (Malas et al., 2003).

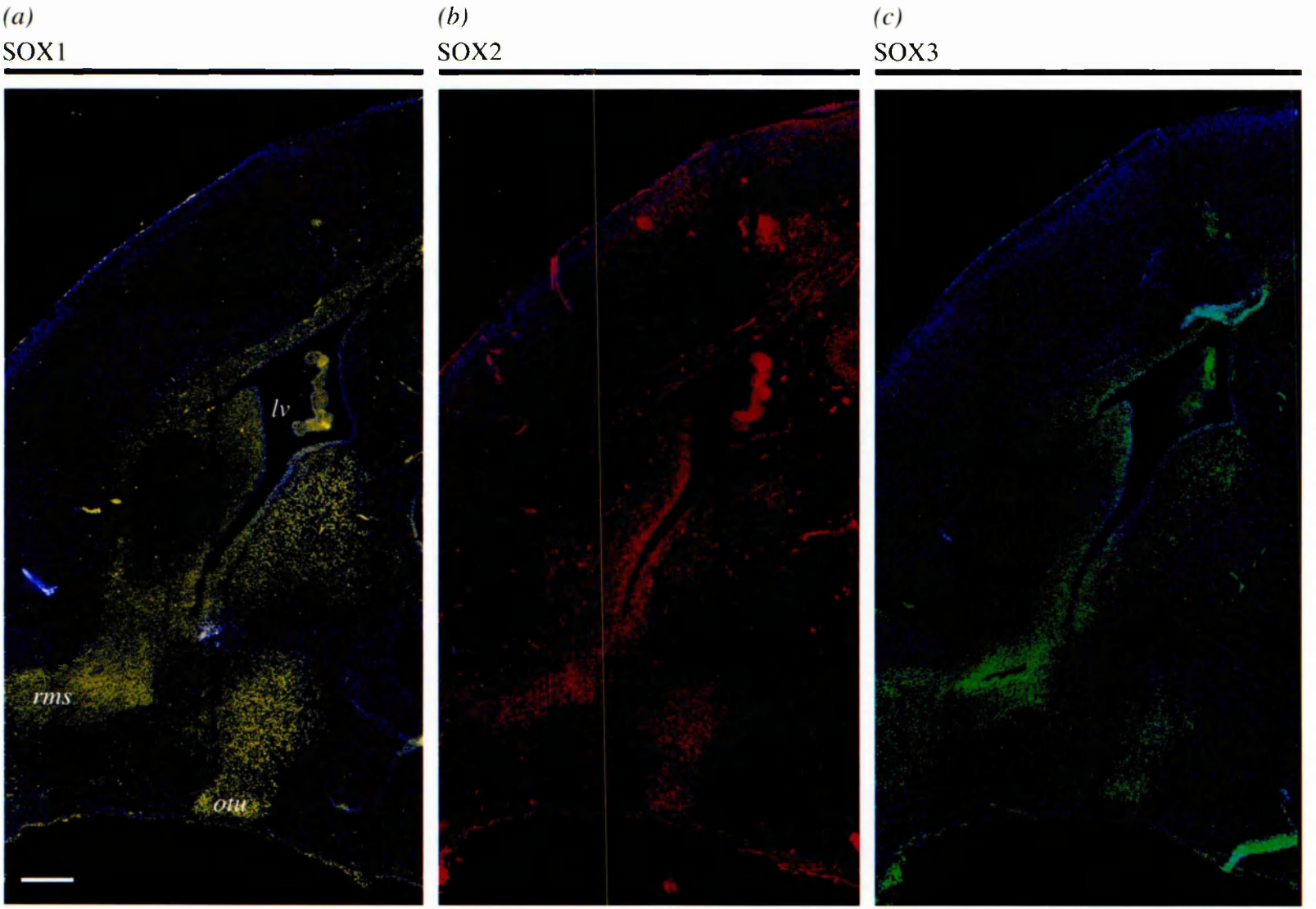
2.3.3 *Sox2* expression in gut and lung.

Staining in the endodermally derived epithelial tissue of the lung, gut and oropharyangeal area is clearly seen in X-gal stained embryos up to 13.5dpc (Fig. 2.2a-h &2.1b(green)). Immunohistochemistry confirms this pattern and reveals that SOX2 protein is localised to the nucleus. SOX2 is detected from the earliest stages

Figure 2.3 SOXB1 expression in the forbrain of 16.5 dpc embryos.

Serial sagittal sections of the anterior portion of 16.5dpc mouse embryo brain. SOX1, SOX2, and SOX3 proteins were detected using specific polyclonal antibodies (materials and methods). Bar is 200μm.

The staining pattern reveals a very similar expression pattern in these structures with cells from the same populations positive for each antibody. New cells are 'born' in the regions surrounding the lateral ventricle (lv) and then migrate away from this site to populate other regions of the developing brain. Much of this migration is radial (such as for the cortex) where as other specific structures are populated by streams of migration of relatively undifferentiated cells such as in the rostral migratory stream (rms) that leads to the olfactory bulbs and migrating cells of the lateral ganglionic eminence that populate the olfactory tubercle (otu). Anterior is to the left.


(a)
SOX1
(b)
SOX2
(c)
SOX3
lv
rms
otu

examined but as the lung develops only the larger of the branching vessels remain positive (Fig. 2.4g,h). Samples of adult lung stained for *ßgeo* activity show that, even at these stages, there continues to be a strong signal in the larger bronchial vessels. Strong staining is seen in the oesophagus and stomach (Fig. 2.4i,k) reiterating the *ßgal* expression observed in $Sox2^{\beta geo-/+}$ embryos (Fig. 2.2). SOX2 is expressed in the oropharyangeal epithelium, including the invagination that forms Rathke's pouch (Kamachi et al., 1998). Patches of staining seen on the tongue of 13.5dpc embryos stained with X-gal indicate that *Sox2* is likely to be expressed in the clusters of cells that will eventually form the taste buds (fungiform papillae) (Fig. 2.4l).

2.3.4 *Sox2* expression in sensory epithelial structures

Sox2 expression has been described as marking the earliest cell populations in sense organ development (Collignon, 1992; Wood and Episkopou, 1999). From 9.5dpc, whole mount staining is unable to clearly show the extent of *ßgeo* expression, but reconstruction of the X-gal pattern from serial sections with pseudo colouring gives a much better indication as to the extent of staining in the developing sense organs (Fig. 2.1b). X-gal staining clearly marks the first stages of eye development with expression at this stage marking optic cup and prospective lens tissue. The expression of *SoxB1* genes has already been described in some detail in the developing eye (Collignon, 1992; Kamachi et al., 1998; Nishiguchi et al., 1998) and expression of *ßgeo* in $Sox2^{\beta geo-/+}$ embryos closely matches this pattern.

At 9.5dpc the olfactory epithelium covers the most anterior aspect of the embryo and $Sox2^{\beta geo-/+}$ samples show *ßgeo* expression throughout this area; a pattern again matched by mRNA *in situ* hybridisation results (see fig 7a and 7b in

Figure 2.4 SOX2 expression in embryonic mouse sections.

SOX2 expression in various sections taken from mouse embryos between 10.5 and 13.5 dpc.

(a). *Transverse section through the head region of a 10.5dpc embryo showing staining in neural epithelium of the CNS, optic cup (**o**), neural epithelium (**nh**) of hindbrain (sorry about the bubble). Bar is 100μm.*

(b). *Transverse section through hindbrain region of an 11.5dpc embryo showing staining in the ependymal layer (**el**) but not the post-mitotic mantle layer (**ml**). Bar is 100μm.*

(c). *Sagittal section through ear of a 13.5dpc embryo showing staining in thickened epithelium of the prospective sensory patches, saccule (**s**), vestibulocochlea VIII ganglion (**g**). Bar is 50μm. **f** Close up of saccule staining in **c**. Bar is 20μm.*

(d). *Sagittal section of 13.5dpc embryo showing the cytoplasmic staining of dorsal root ganglia cells. Bar is 100μm. **e** Close up of **d**. Bar is 20μm.*

(g). *Sagittal section through the lung of a 13.5dpc embryo showing staining in the epithelial layer of bronchus (**b**). Bar is 100μm. **h** Close up of **g**. Bar is 20μm.*

(i). *Transverse section at the mouth level of an 11.5dpc embryo showing nuclear staining of the epithelium of the pharangeal region of the foregut (**f**) and the laryngeal orifice (**lo**). Bar is 20μm.*

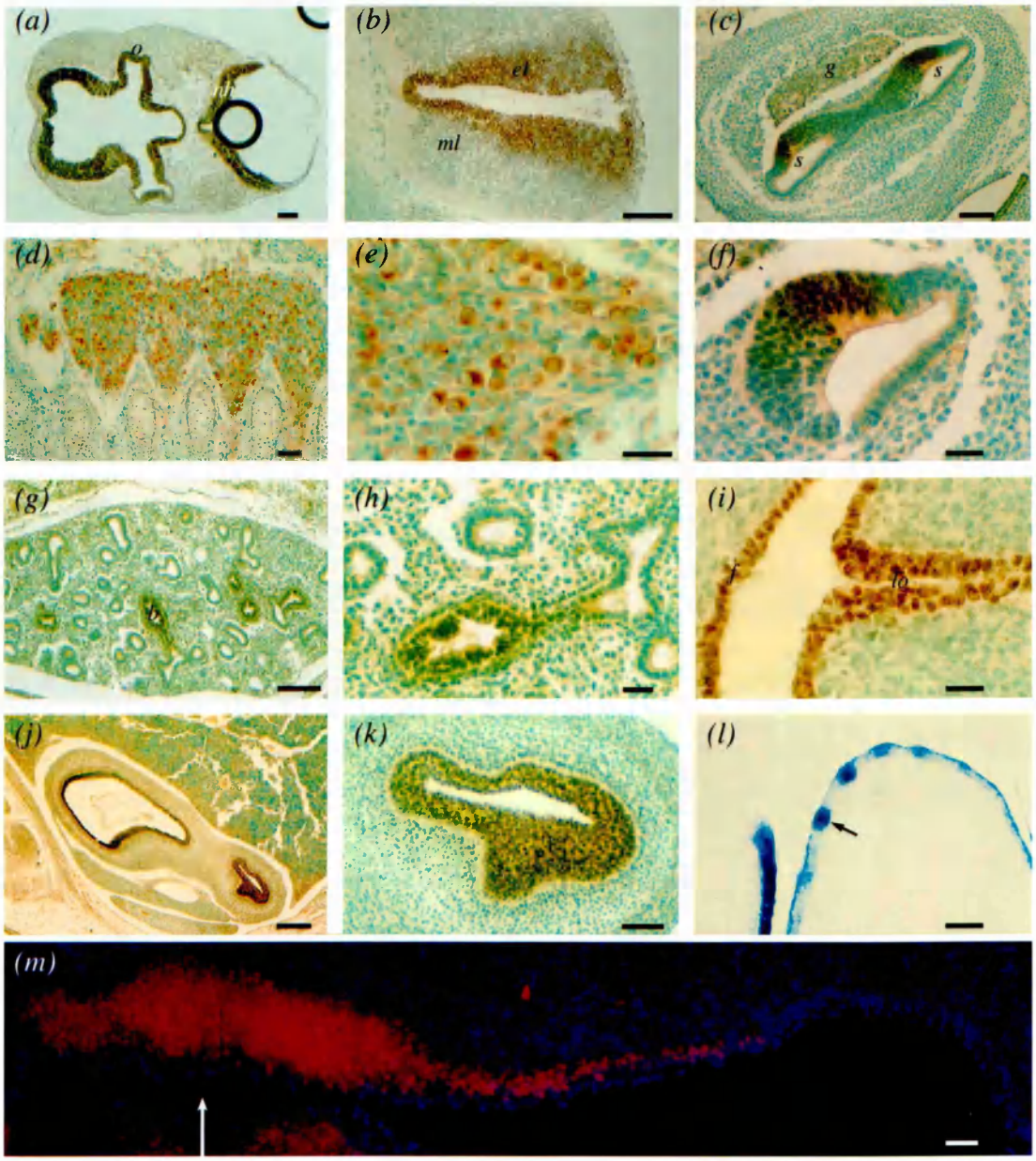
(j). *Sagittal section through stomach of a 13.5dpc embryo showing staining in the epithelial lining. Bar is 200μm. **k** close up of **j**. Bar is 50μm.*

(l). *Sagittal section of 13.5dpc tongue showing epithelial staining, particularly in the fungiform papillae (arrow). Bar is 50μm.*

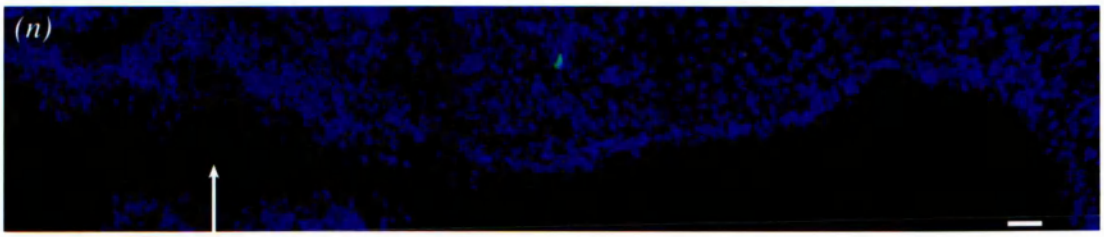
(m). *Fluroescence immunohistochemistry to detect SOX2 (red) in the olfactory epithelium of a 14.5dpc embryo. This section was then subject to SOX3 immunohistochemistry shown in n. Arrow marks midline. Bar is 20μm.*

(n). *Fluorescence immunohistochemistry showing that SOX3 (green) is not present in the olfactory epithelium. This is the same section as m. Sequential immunohistochemistry was performed as described in materials and methods. As an internal control SOX3 was detected in parts of the CNS on this section (data not shown). Arrow marks midline. Bar is 20μm.*

SOX2



SOX3



(Collignon, 1992; Kamachi et al., 1998)). Wholemount X-gal staining of *Sox2*$^{\beta geo-/+}$ embryos indicates the continued expression of *Sox2* in the nasal epithelium (Fig. 2.2a-h). By 14.5dpc the SOX2 antibody detects protein in the medial portion of the epithelium lining the nasal cavity whilst no SOX3 (or SOX1) can be detected at this stage (Fig. 2.4m,n) despite earlier overlapping expression in these endodermally derived tissues ((Collignon, 1992; Wood and Episkopou, 1999) and Brunelli unpublished data).

Sox2 expression in the developing Ear

Sox2 is detected during the formation of the embryonic ear from the earliest stages of otic placode formation (Collignon, 1992; Wood and Episkopou, 1999). X-gal staining of the developing ear in *Sox2*$^{\beta geo-/+}$ embryos matches that observed with the antibody and shows a pattern marking the prospective sensory epithelia that becomes increasingly restricted as this complex structure develops.

The ear is an elegantly complex organ responsible for the senses of hearing and of motion and position. The molecular mechanisms dictating the morphogenesis and differentiation of the mammalian ear are largely unknown although a number of genes have been linked to both syndromic and non-syndromic deafness in humans (Steel and Kros, 2001). Further genes have been implicated in various aspects of ear development through null mutations and expression studies in mouse, chick and frog (Torres and Giraldez, 1998).

Sensory cells responsible for detecting sound or motion and transmitting them to the brain are located in the inner most part of the ear structure. Mammals have a

Figure 2.5 *X-gal* staining of *Sox2*$^{\beta geo-/+}$ embryos marks *Sox2* expression in the inner ear.

(a). *Whole mount X-gal staining of a 11.5 dpc* Sox2$^{\beta geo-/+}$ *embryo. The outline of the otic vesicle is difficult to visualise against the background of staining in the CNS but is marked in the high magnification view where some blue staining can be seen. For orientation purposes D is dorsal, L is lateral and A is anterior. Bar is 200μm.*

(b). *After wax embedding and sectioning staining of the otocyst can be clearly seen in the thickened epithelium of the developing inner ear. Computer reconstruction of from sections shows how staining relates to the entire structure. Example sections below correspond approximately to the regions marked on the model. Light blue corresponds to space within the otocyst. Green, yellow and red mark low to high intensities of X-gal staining respectively. Reconstruction of this structure from sections shows how the ßgeo expressing cells are wrapped around the lumen of the developing inner ear and mark regions where sensory epithelium will develop in both the prospective cochlea and vestibular components. Bar is 200μm.*

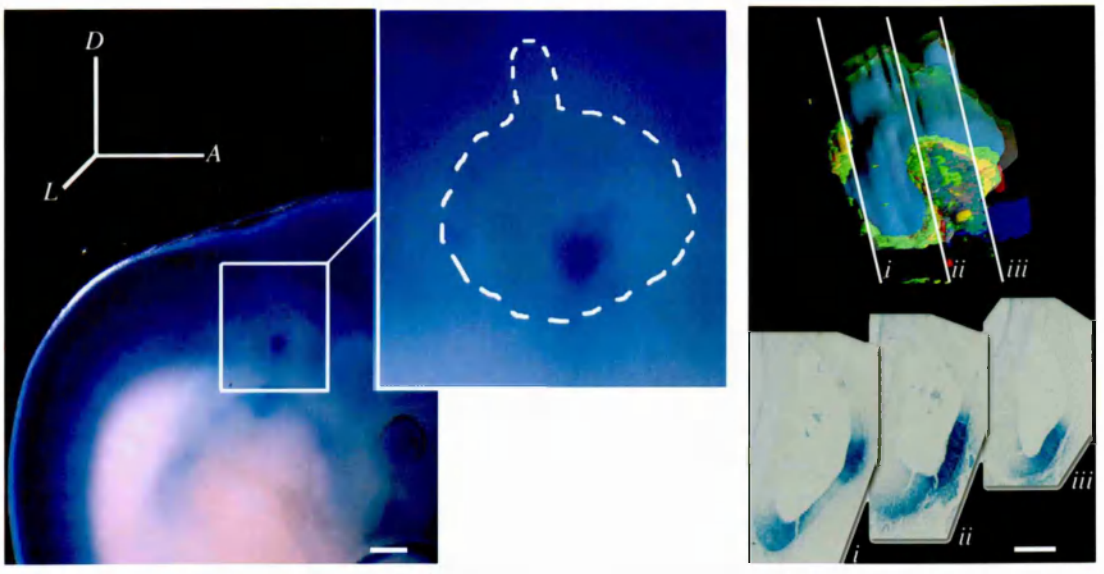
(c). *Computer reconstruction of the developing inner ear of Sox2*$^{\beta geo-/+}$ *embryo at 9.5dpc showing X-gal staining. Yellow and red colour marks X-gal staining with transparent grey indicating the shape of the otic epithelium. Bar is 100μm.*

(d), (e). *Similar to* ***(c)*** *but with the inner ear of a 10.5dpc and 12.5dpc Sox2*$^{\beta geo-/+}$ *embryo shown respectively. Bars are 100μm.*

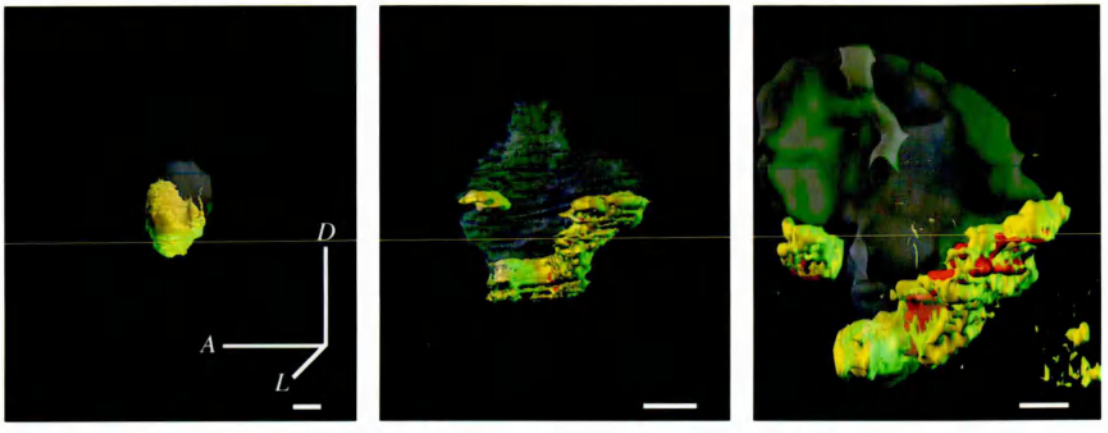
(f). *Immunohistochemistry using a SOX2 specific antibody marks the prospective sensory epithelial cells of the cristae in a 13.5dpc mouse embryo. Bar is 20μm.*

(g). *Schematic representation of the inner ear structure at 14.5dpc showing the localisation of sensory epithelia. cO cochlea, c cristae, mu macula utriculi, ms macula sacculae.* ***(Fekete, 1999).***

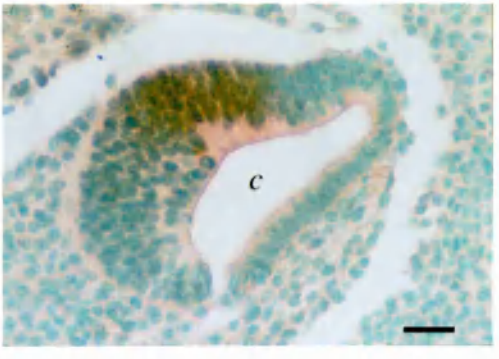
(h-j). *Serial section immunohistochemistry using antibodies specific for* SOX1 ***(h),*** SOX2 ***(i)*** *and* SOX3 ***(j)*** *only detects the presence of* SOX2 *protein in the mouse inner ear at 14.5dpc. Staining is observed in the prospective sensory epithelia of both the cochlea and vestibular halves of this structure. Success of the immunohistochemical procedure where no inner ear staining was detected was determined by observing appropriate staining in other structures on the same section (data not shown). Labels as for* ***(g).*** *Bars are 100μm.*



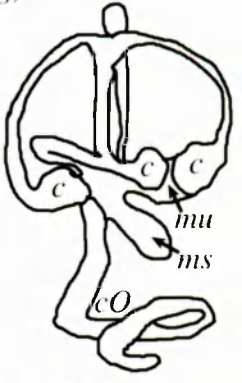
(c) (d) (e)



(f) SOX2 (g)

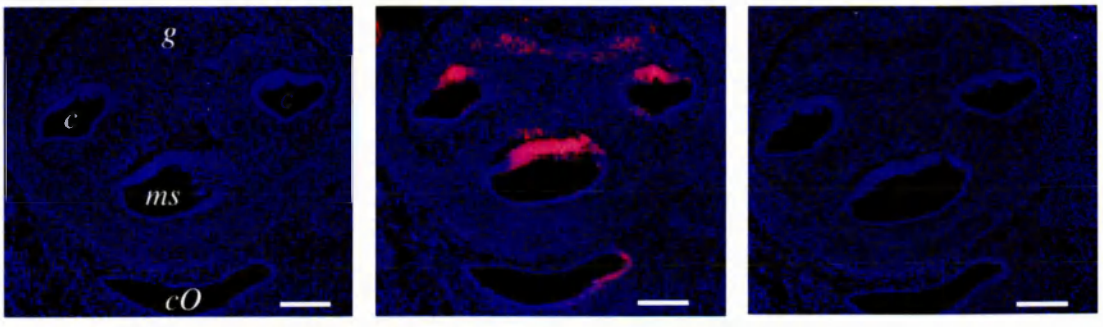


(h) SOX1



(j) SOX3

(i) SOX2



pair of ears and each of these derives from an area of epithelium adjacent to the neural epithelium of the headfold. By 9.5dpc these otic placodes invaginate forming hollow epithelial spheres, the otic vesicles, lateral to rhombomere four of the hindbrain. Almost all of the cellular components of the inner ear derive from the otic vesicle although there is a small contribution from a neural crest derived melanocyte population (Roydon-Price and Fisher, 2001). From a simple structure at this stage cells divide and differentiate to produce a highly complex and organised form capable of sensing sound as well as changes in position in space. Both auditory and motion senses rely upon the endolymphatic fluid which fills the inner ear. External sound or change in aspect causes movement of this fluid, stimulating hair cells of the respective sensory epithelia. The cochlea is the component of the inner ear that is responsible for the detection of sound whereas the vestibular sensory cells of the macula sacculae, macula utriculi and the cristae record changes in gravity or angular acceleration. Sound waves collected through the outer ear are transduced and amplified mechanically through the middle ear and are finally converted into vibrations of the endolymph of the inner ear. The spiral shaped cochlea has an array of auditory hair cells aligned along its length, collectively called the organ of corti. Different amplitudes and frequencies of vibration will stimulate particular sets of hair cells in the cochlea to different extents and in this way a wide range of sounds can be detected and then interpreted by the brain.

Three-dimensional reconstruction from serial sections of these embryonic inner ears make it easier to visualise areas that express *βgeo* in the initial stages of otocyst formation, between 9.5 and 12.5dpc (Fig. 2.5a-e). This expression pattern is confirmed by anti-SOX2 immunohistochemistry (Fig. 2.5f,i). Models and section

data indicate that *Sox2* is expressed in all the cells from which the sensory epithelium derives as well as in the ganglia that innervate them (Fig. 2.5).

It has already been reported that *Sox2* marks the prospective otic placodes at 8.5dpc, but at 9.5dpc *Sox2* is detected only in the ventral half of the otic vesicle (Chapter 1) (Fig. 2.5c). By 11.5dpc the otic vesicle has begun to resolve its final structure with a clear division between the auditory and vestibular halves of the inner ear. The 3D model representing the expression pattern of *Sox2* at this stage shows the protein localising to the developing cristae and maculae at the base of the vertical canal plate and extending along the side of the cochlea anlagen that will develop into the organ of corti. At 13.5dpc and 14.5dpc immunohistochemistry confirms that SOX2 is present only in the prospective sensory cells of the vestibular and auditory components of the inner ear and that both SOX1 and SOX3 are absent (Fig. 2.5h-j). Staining observed in thickened regions of the otic epithelia confirms that they are likely to be prospective sensory cells. A signal is also detected in cells of the facioacoustic (VII-VIII) ganglion complex that will eventually innervate the fully formed inner ear (Fig. 2.5i).

2.3.5 *Sox2* expression in the hair follicles of the vibrissae and pelage.

A site of *Sox2* expression that has not been described previously is the developing hair follicle. *Sox2* mRNA expression is detectable at 13.5dpc just under the epithelial layer adjacent to the nostrils. Radioactive mRNA *in situ hybridisation* indicates that there is expression of *Sox2* in some component of the developing vibrissae but because this procedure has no cellular resolution it is impossible to accurately ascertain exactly which cells these are (Fig. 2.9h) (Collignon). Various

stages of follicle development were therefore analysed in order to clarify the expression pattern of *Sox2* and to elucidate how it might be contributing to follicle development.

The hair follicle is just one of a number of epithelial appendages in diverse species that require a series of reciprocal inductive interactions between the epithelium and the mesenchyme for their correct formation, the first steps of which appear to be evolutionarily conserved. In mice four discrete waves of follicle induction occur, each giving rise to a particular type of hair but the general processes involved in induction of all hair follicles appear to be the same (Hardy, 1969; Slee, 1962). Two main phases are involved in follicle induction, initially the number and spacing of placodes is determined and then this is followed by cellular differentiation that facilitates the formation of one of a myriad of possible structures.

Follicle induction

The first of these signals defines areas that are competent to form appendages and tissue recombination experiments demonstrate that they derive from the mesenchyme (Fig. 2.6a) (Hardy, 1992; Oro and Scott, 1998; Reddy et al., 2001). Focal points of appendage formation are defined, possibly through a reaction-diffusion mechanism (Slack, 1991), and cause regions of epithelium to thicken forming placodes. From these points a second, epidermally derived, signal instructs a small group of underlying mesenchyme cells to condense (Fig. 2.6b). From this group of cells the third message instructs epithelial differentiation and proliferation to produce an appendage appropriate to its class of origin, whether that be scale, feather

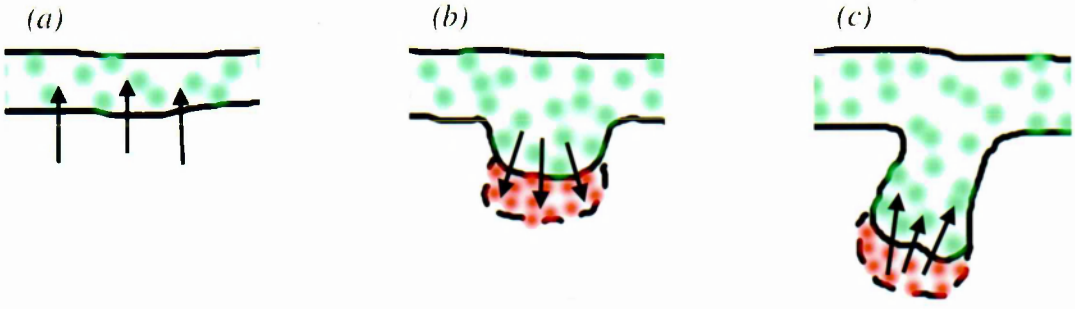
Figure 2.6 Epithelial-mesenchymal interactions during follicle development.

Figure adapted from M.H.Hardy (1992) showing the flow of signals in the early stages of follicle development.

(a). *First , mesenchymal, signal indicating 'make an appendage'.*

(b). *Second, epithelial, signal defining point at which appendage is to be made.*

(c). *Third, mesenchymal, signal directing what sort of appendage to make.*


(a)
(b)
(c)

or hair (Fig. 2.6c). It has become clear that there is a range of factors and collaborations involved in the formation of these epithelial appendages that may throw light on developmental processes common to other sites of organogenesis.

Pathways involved in the signalling interplay between mesodermal and epidermally derived structures appear to be complex but a few genes have been shown to play important roles in this process. The initial patterning of placodal fields appears to rely on the interplay between several factors such as Wnts, FGFs, BMPs, Notch and Delta and Noggin (Millar, 2002).

WNT signalling and platelet derived growth factor-A (PDGF-A) contribute to the second (epithelial) signal triggering the induction of the dermal condensate. *Sonic hedgehog* (*Shh*) expression depends upon the WNT signal and is required for the development of the dermal condensate into the dermal papilla. In *Shh* deficient mice there is no down growth of the follicular epithelium and so it is likely to be involved in initiating the third signal from the dermal condensate which might involve the production of activins/BMPs and their inhibitors (Millar, 2002).

Follicle growth

The progression of epithelial proliferation in hair development produces a plug that elongates into the mesodermal layer. The position of the mesodermal condensation is maintained at the base of the plug as two layers of epithelial cells form the inner and outer root sheath (IRS and ORS). Subsequent proliferation almost completely surrounds the dermal condensate, which develops into the dermal papilla, with cells adjacent to this structure forming the matrix. Matrix cells proliferate transiently and move toward the surface with central matrix becoming pre-cortical

cells that subsequently give rise to the cortex, medulla and cuticle of the hair shaft. More peripheral matrix cells form the IRS which is in turn surrounded by ORS. Near the skin surface the IRS degenerates, allowing the newly formed hair shaft to be pushed outward as further matrix cells differentiate (Fig. 2.7).

Follicle cycling

Once fully formed, follicles enter a cycle of active growth (anagen) regression (catagen) and rest (telogen) that is normally maintained throughout postnatal life. It is only the lower half of the hair follicle that undergoes change, with cells for follicle renewal residing in the bulge region at the base of the permanent portion. At the end of anagen matrix cells lose their proliferative capacity and the lower part of the follicle regresses bringing the dermal papilla closer to the bulge region (Fig. 2.7 *Catagen*). At the telogen/anagen transition a signal, which is perhaps from the dermal papilla, triggers proliferation of a few stem cells that produce new matrix and ORS that differentiate and move downward(Fig. 2.7 *Anagen I*). When the lower half of the follicle is reformed once more, further proliferation becomes restricted to the matrix cells that are in contact with the dermal papilla as the follicle progresses through a new cycle (Fig. 2.7 *Anagen IV-VI*).

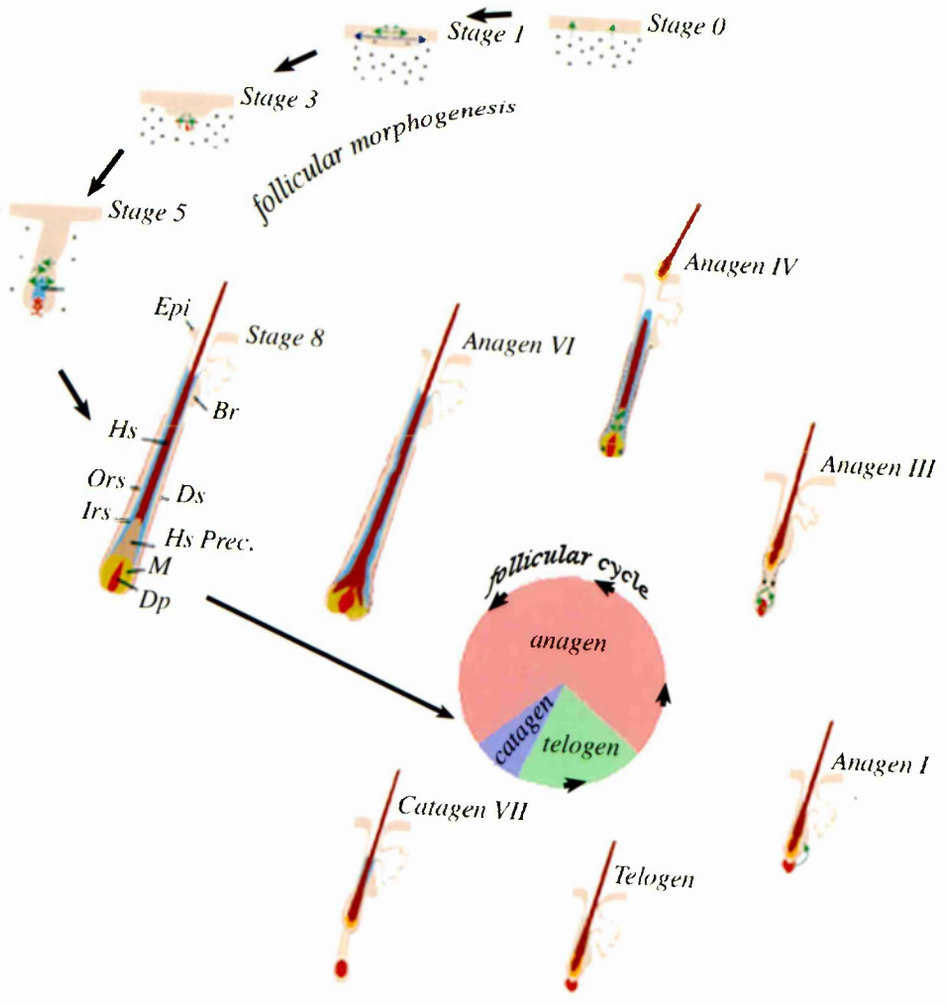
2.3.6 Whole mount X-gal staining of the developing hair follicles in *Sox2*$^{\beta geo-/+}$ embryos.

From 12dpc small spots of X-gal staining in the hair germs of the vibrissae begin to appear in *Sox2*$^{\beta geo-/+}$ embryos. This coincides with the first morphological signs of induction of these follicles (Fig. 2.2d). Vibrissae form in rows lateral to the nostrils and just a few hours after the first induction is observed the X-gal staining

Figure 2.7 The stages of hair follicle formation and cycling.

Cartoon indicating the key stages in hair follicle induction and cycling. Morphogenesis occurs only once in the lifetime of an animal but the cycle of growth, regression and moulting occurs many times over.

Epithelium (Epi), Bulge region (Br), Hair shaft (Hs), Dermal sheath (Ds), Inner root sheath (Irs), Hair shaft precursors (Hs Prec.), Matrix (M), Dermal papilla (Dp).


Stage 1
Stage 0
Stage 3
follicular morphogenesis
Stage 5
Anagen IV
Epi
Stage 8
Anagen VI
Br
Hs
Anagen III
Ors
Ds
Irs
Hs Prec.
M
Dp
follicular cycle
anagen
catagen
telogen
Anagen I
Catagen VII
Telogen

pattern becomes more extensive as the follicles develop and additional populations of cells in each begin to express *ßgeo*. At 13.5dpc the induction of this hair type is complete and X-gal staining marks all the vibrissae formed including those above the eye and on the cheek (Fig. 2.2g,h). Also at 13.5dpc is the first appearance of the pelage follicles with a few spots of *ßgeo* expression detectable over the lateral aspect of the abdomen (Fig. 2.2g). By 15.5dpc X-gal reveals a fairly evenly spaced, uniformly stained set of spots over most of the surface of the embryo. This corresponds to the completion of the first wave of pelage follicle induction (Fig. 2.8a). Follicles induced in this round of induction will produce monotrick hairs that are the longest hair type in the adult coat (Dry, 1926; Hardy, 1969). At 17.5dpc extensive staining of the large whisker follicles is observed and a similar, although smaller, version of this staining is seen in the monotrick pelage follicles. Shortly after the primary spate of pelage follicle formation a second wave of induction produces the placodes of the outer coat that contains the hair types awls and auchines. Once again the *ßgeo* spots appear in a roughly hexagonal array in between the sites of monotrick induction (Fig. 2.8d). Just before birth, there have been two rounds of pelage hair follicle induction in the skin and it is clear that all the follicles formed by this point show X-gal staining in the $Sox2^{\beta geo+/-}$ embryos. By 17.5dpc strong blue staining of the entire dermal sheath of the developing vibrissae give a "comma" like appearance to these well established hair follicles (Fig. 2.8d).

Figure 2.8 X-gal staining of *Sox2*$^{\beta geo-/+}$ embryos marks *Sox2* expression in the hair follicles.

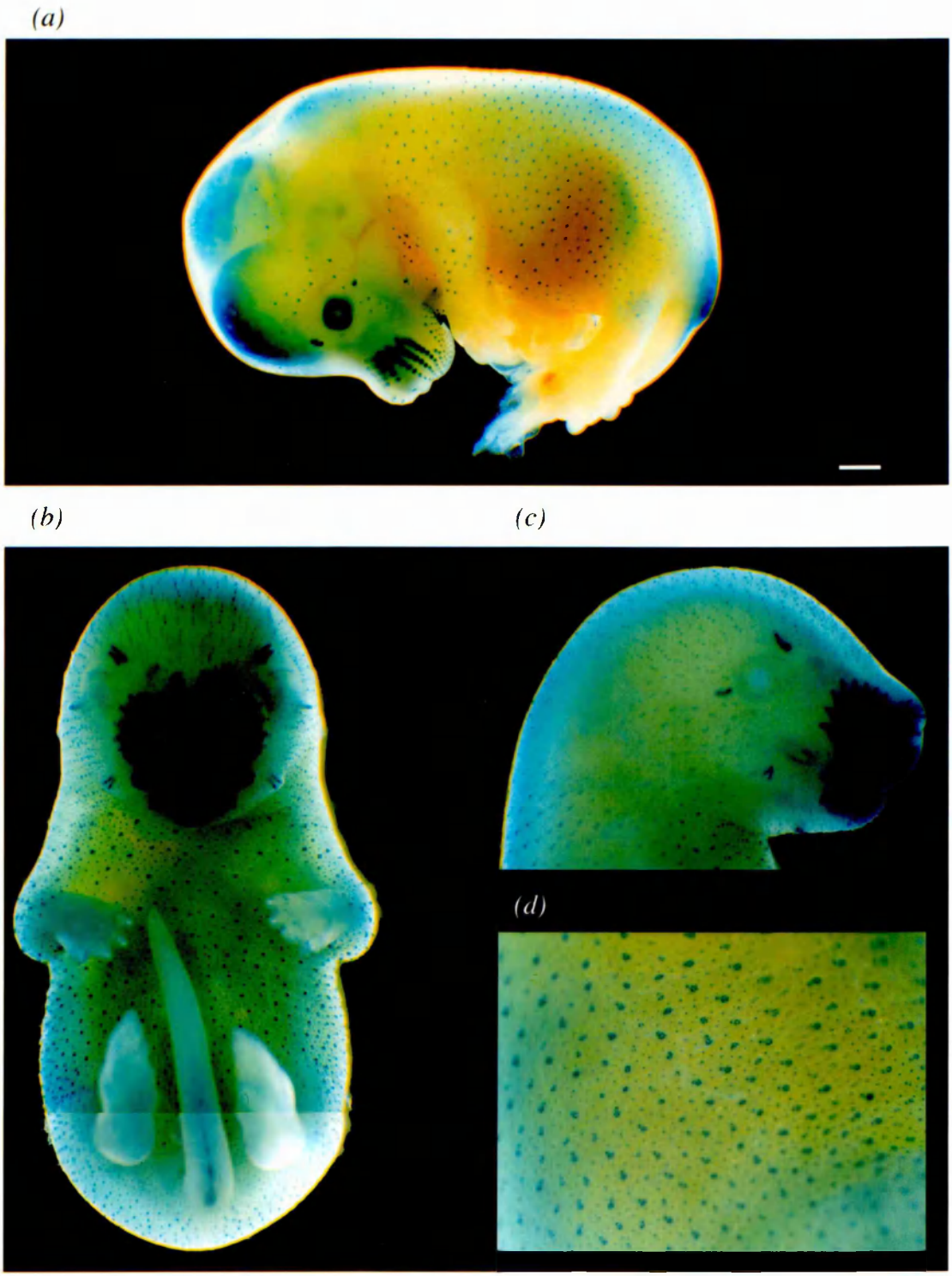
(a). *By 15.5dpc the first wave of pelage hair follicle induction is almost complete. ßgal expression reveals a fairly evenly spaced pattern of dermal papilla that belong to follicles that will give rise to the largest hairs in the mouse pelage, the guard hairs or monotricks.*

(b). *At 17.5dpc two populations of pelage follicles can be seen, the larger spots mark those ones developed from the first wave of induction with smaller spots indicating follicles formed in a second wave.*

(c). *Higher magnification view of the head of the embryo in* ***b*** *showing the lines of vibrissal follicles that have formed lateral to the nostrils. These follicles are no longer marked by just a spot of expression but have many cells staining blue.*

(d). *Higher magnification view of the 17.5dpc skin showing follicles induced over two rounds. The larger first round monotrick follicles have a similar expression pattern to those of the vibrissae but are smaller.*

Bar is 1mm for ***a,b*** *and* ***c****, and 500µm for* ***d****.*


(a)
(b)
(c)
(d)

2.3.7 Sectioning reveals the detail of *Sox2* expression pattern in individual hair follicles

The staining pattern of individual follicles broadens as they develop and this is highlighted in histological sections of X-gal stained $Sox2^{\beta geo-/+}$ samples and by section immunohistochemistry with a SOX2 antibody.

SOX2 immunohistochemistry confirms that the protein appears at the initial stages of hair induction in both vibrissae and pelage follicles in an identical pattern to that of *ßgeo* in $Sox2^{\beta geo-/+}$ embryos (Fig. 2.9a-f). SOX2 is detected from the earliest stages of follicle development in the nuclei of cells that have responded to the second inductive signal. All dermal condensations are marked by SOX2 expression soon after the adjacent epithelial placodes begin to thicken (Fig. 2.9a,d). This cluster of cells maintains SOX2 expression as the epithelial component of the follicle proliferates forming the first inner and outer root sheath cells (Fig. 2.9b,e). Once created the dermal condensation and subsequently the dermal papilla is believed to direct the overlying epithelial cells to form the highly organised follicle structure containing six concentric cylinders of cells each synthesising its own distinctive set of proteins. More developed follicles display further staining around the invaginating follicular epithelium. These are the mesodermally derived dermal sheath that show X-gal (Fig. 2.9k) as well as immunohistochemical staining for SOX2 (data not shown). They have many dermal papilla properties and have been proposed to repopulate it throughout the life of the animal (Reynolds et al., 1999).

Further characteristic expression of SOX2 appears to be limited to vibrissae and the first follicles to develop on the torso that give rise to monotrichs. In these follicles

Figure 2.9 Sox2 *expression in developing hair follicles.*

This series of images shows how the expression of SOX2 closely matches that of ßgeo expression from the earliest stages of hair follicle development. Bar is 20μm in all pictures.

(a)-(c). *X-Gal staining of early stage hair follicles in* $Sox2^{\beta geo+/-}$ *embryos.* ***a*** *stage 0-1,* ***b*** *stage 2,* ***c*** *stage 3a.*

(d)-(f). *SOX2 localization in early stage hair follicles, following staining in the condensed mesenchyme (arrow head). Occasional staining of epidermal cells can be detected in later stage follicles. (arrow).* ***d*** *stage 0-1,* ***e*** *stage 2,* ***f*** *stage 3a.*

(g). *Representative pre-immune serum control. This particular serial section equivalent to* ***e***

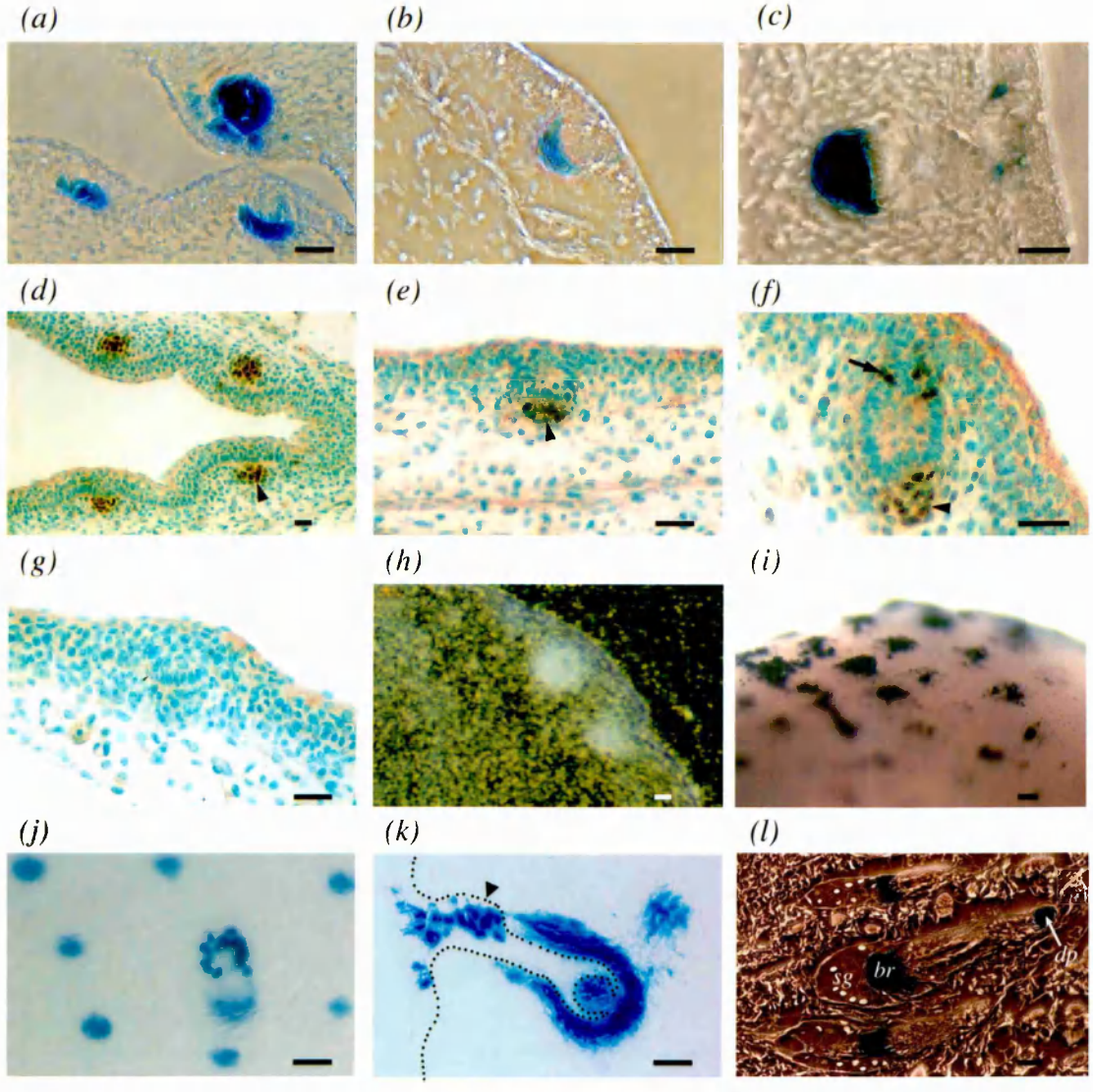
(h). *Detection of* Sox2 *RNA in 13.5dpc vibrissae by radioactive* in situ *hybridization (J.Collignon).*

(i). *Detection of Sox2 RNA in the developing feather buds of stage 34 chick skin.*

(j). *X-gal staining in the skin of a 16.5dpc* $Sox2^{\beta geo+/}$ *embryo. The second component of staining is seen in the more mature follicle (centre). Note the similarity with figure F.*

(k). *A stage 5 vibrissal follicle which indicates the three components to SOX2 expression in this structure. The dermal papilla (dp), the dermal sheath (ds) and the epithelial cells that may be Merkel cells (br) these reside within the 'bulge region' of the hair follicle (arrow head). The dotted line marks the boundary between the epidermis and dermis.*

(l). *X-gal stained adult hair follicles. In this section a catagen phase hair follicle shows staining of the dermal papilla (dp) and a cluster of cells, just below characteristic cells of the sebaceous gland (sg), that is referred to as the bulge region (br).*


(a)
(b)
(c)
(d)
(e)
(f)
(g)
(h)
(i)
(j)
(k)
(l)
sg
br
dp

SOX2 positive cells are first detected in the ORS from stage 3c onwards (Fig. 2.9f). In whole mount they are seen as an incomplete ring of cells close to the surface of the skin (Fig. 2.9j). At later stages most stained cells are found in the bulge region with fewer detected in other parts of the isthmus (Fig. 2.9k). This 'horseshoe' pattern is reminiscent of the expression of mouse atonal homologue-1 (MATH1) that marks Merkel cells in the developing hair follicle (Ben-Arie et al., 2000). Merkel cells are mesodermally derived slowly adapting mechanoreceptors that are activated by steady skin indentation (Szeder et al., 2003). Stem cells, that support the cyclical growth and regression of the fully formed hair follicle, are also thought to reside in the bulge region and their behaviour may be influenced by the Merkel cells (Moll et al., 1996; Oshima et al., 2001).

2.3.8 *Sox2* expression in adult (cycling) follicles.

All three components of the staining pattern persist in anaphase monotrick or vibrissae hair follicles from adult animals. Regressing or resting stage follicles appear to only maintain *Sox2* expression in the dermal papilla and in cells in the bulge region that lies next to the characteristically large cells of the sebaceous gland (Fig. 2.9l). No expression of the most closely related *SoxB1* genes *Sox1* and *Sox3* was detected in any of the stages of follicle development.

2.3.9 Expression of *Sox2* in feather buds in chick embryos.

The first stages of feather formation are very similar to that of hair with interaction between cell layers forming a epithelial placode and an underlying mesodermal condensation. Cell proliferation as well as recruitment of mesodermal cells produces a feather bud that first protrudes out of the skin surface before

invaginating inward to form a feather follicle. The mesodermal component forms the dermal papilla and is surrounded by the collar, which equates to the matrix region of a hair follicle, and is the region where new epithelial cells are added to the growing feather. The characteristic shape of the mature feather is achieved by complex combination of epithelial invagination, fusion and cell death that gives rise to feather barbs and barbules. *Sox2* is expressed in murine epithelial appendages but it is unclear as to whether this is a conserved site of expression in diverse species. To test this, mRNA *in situ* hybridisation was performed on the initial stages of feather bud formation in chick embryos with preliminary analysis indicating the presence of *Sox2* transcripts (Fig. 2.9i).

The signal is first detected in patches of dorsal skin of stage 33 chick embryos (Hamburger and Hamilton, 1951) just before there is any morphological sign of feather bud induction (Jung et al., 1998). As feather bud development progresses the signal becomes slightly more widespread with *Sox2* appearing to be expressed in all the mesodermal cells within the protruding feather bud. The results show that in the initial stages of epithelial appendage formation *Sox2* is expressed in equivalent structures in chick and mouse and so may be an evolutionarily conserved site of expression.

2.4 Discussion

The comprehensive examination of expression patterns provides invaluable information with regard to understanding the role a particular gene has to play during development. An expression pattern may mark seemingly diverse sites that are linked by a particular property. Expression data could therefore provide clues to the role and possibly mechanism of action of any given gene.

Previous examination of *SoxB* gene expression including *Sox2* has focussed on stages prior to midgestation with the exception of a few specific sites (Chapter 1). During these stages *Sox2* expression is generally limited to relatively undifferentiated cells such as the epiblast, neuroepithelium of the CNS and gut endoderm. A further examination of *Sox2* was undertaken primarily to investigate the possibility of uncategorised sites of expression in the second half of embryogenesis and to identify sites of expression that might be interesting to study by the conditional null mutation of *Sox2*.

This investigation also provided the opportunity to describe *Sox2* expression at later developmental stages and confirmed that the expression of *βgeo,* from the targeted allele $Sox2^{\beta geo}$, is a reliable indicator for the expression of *Sox2*. Having established this, $Sox2^{\beta geo-/+}$ tissue may be used to isolate pure populations of *Sox2* expressing cells for *in vitro* experimentation, either by antibiotic selection or by fluorescence activated cell sorting (Li et al., 1998; Zappone et al., 2000). One use for *Sox2* expressing cells from diverse sites may be in identifying common targets or co-regulators by comparative analysis of expressed genes. In this instance it is important

to know exactly which cells are expressing the selectable marker so that particular populations may be isolated.

After midgestation SOX2 expression is maintained in many of the tissues where it is first detected at earlier stages in development (Chapter1), (Wood and Episkopou, 1999). In the CNS *Sox2* continues to be expressed predominantly in sites of neurogenesis for example in the ventricular and sub-ventricular zones of the developing forebrain. In these areas *Sox2* shares expression with its closest relatives *Sox1* and *Sox3*. SOX2 is also detected, along with SOX1 and SOX3, in many of the cells contributing to the olfactory tubercle although SOX1 staining does appear a little more extensive in this region. The null mutation of *Sox1* produces a severe disruption of olfactory tubercle development, but there is no obvious consequence to its absence in other areas of the developing CNS (Chapter 1) (Malas et al., 2003). This suggests that functional similarities with the other *SoxB1* genes may be sufficient to compensate for the loss of SOX1 in many cells where there is overlapping expression, but there may be some unique role for SOX1 in olfactory tubercle development. Similar unique roles may be revealed upon conditional disruption of *Sox2* or *Sox3* in the developing CNS. Where there is overlapping expression and functional compensation, multiple null alleles may have to be introduced before phenotypes are exhibited and the role of these genes in these cells deciphered (Chapter 5).

Sox2 has a broader expression pattern than either *Sox1* or *Sox3* with the most extensive sites of overlapping expression limited to the developing CNS. At early stages there is also overlapping expression of *Sox2* and *Sox3* in the sensory placodes

of the nose and ear but by 13.5dpc *Sox3* is no longer co-expressed in these tissues where *Sox2* only is observed (Fig. 2.4m,n & Fig. 2.5i,j). In the developing lens *Sox1* and *Sox2* initially have overlapping expression but *Sox2* is down regulated so that by 12.5dpc only *Sox1* is expressed. This is the same stage at which lens abnormalities appear in $Sox1^{m1-/-}$ animals indicating that in the lens, up until this stage, functional compensation by *Sox2* is sufficient for normal formation (Nishiguchi et al., 1998). Similar compensation may also occur at other sites of overlapping expression. If this is the case then null mutation may be expected to cause the most severe phenotypes in tissues where there is unique expression or where overlapping expression of close relatives has ceased. This information may be critically important when attempting to link the phenotypes of uncategorised mouse mutants with a specific gene.

The initial stages of inner ear formation are marked by overlapping expression of *Sox2* and *Sox3,* but *Sox3* expression is soon down regulated leaving *Sox2* as the only *SoxB1* gene expressed in this region (Fig. 2.5i,j) (Uchikawa et al., 1999; Wood and Episkopou, 1999). *Sox2* marks sensory placode cells from the outset of inner ear formation (Wood and Episkopou, 1999) and expression in the neuroepithelium persists as the major elements of this complex structure are formed. Three dimensional reconstruction shows *Sox2* in the ventral half of the otic vesicle at 9.5dpc and it persists in cells of the prospective sensory epithelium as the structure of the inner ear develops (Fig. 2.5c-e). By 14.5dpc SOX2, but not SOX1 or SOX3, is detected in the prospective sensory cells along one side of the cochlea and in vestibular neuroepithelium (Fig. 2.5h,i,j).

SoxB1 expression in chick differs slightly in that all three genes are detected during inner ear formation (Uchikawa et al., 1999). *Sox2* and *Sox3* are detected from stage 9-10 (~8.5dpc mouse) in the otic placode whilst *Sox1* appears slightly later, just before stage18 (~9.5dpc mouse). Stage 28 chick embryos (~12.5dpc mouse) have strong and uniform expression of *Sox2* in all the sensory neuroepithelial patches of the inner ear, *Sox1* is strong in the cristae but weak in other sensory patches and *Sox3* is detectable in the cristae only. These observations indicate a degree of regulatory conservation between these two species and differences may simply reflect the alternative evolution of regulatory elements where there was once functional redundancy. Perhaps a more interesting possibility is that these alternative patterns represent the evolutionary fine tuning of an ancestral inner ear structure that has produced differences in the relative levels of these transcription factors in the respective species. Even slight differences in activity may have favoured the conservation of an individual gene's expression during the evolution of an optimal regulatory architecture. This possibility may be investigated to some degree by targeted exchange of the coding region of these genes and is discussed further in chapter 5.

Sox2 expression in vibrissal and pelage hair follicles was described for the first time in this study with staining observed in more than one component of the developing follicle. Detection of SOX2 in precursors of receptor cells correlates well with expression in the other sensory organs, the tongue and perhaps in the lining of the gut and lung. In these locations SOX2 may have some role in determining the identity of specialised cells required for detecting and transducing changes in their environment. Other *Sox2* expressing cells in the hair follicle have a completely

different function. From the earliest stages of hair follicle development SOX2 is detected in dermal condensations that form in response to an epithelial signal, possibly involving the *Wnt* signalling pathway and platelet derived growth factor-A (PDGF-A) (DasGupta and Fuchs, 1999; Millar, 2002). Once initiated, *Sox2* expression is maintained in this group of cells into adulthood with no expression of either *Sox1* or *Sox3* ever detected here. SOX2 is also observed in the dermal papilla during feather formation indicating that this is an evolutionarily conserved site of expression. In both species this group of cells directs the correct formation of overlying epithelium, both initially and during successive rounds of growth and regression. Cross talk between opposing cell layers is a common theme throughout development and epithelial-mesenchymal interactions have been demonstrated to regulate chick *cSox2* in both the developing gut and lung (Ishii et al., 1998). Similar interactions may also be working to pattern these structures in the mouse. The dermal papilla is critical in the complex reciprocal signalling between juxtaposed cell layers and grafting experiments indicate that there is a defined progression of signalling in epithelial appendage development (Fig. 2.6) (Hardy, 1992).

There is also overlapping expression with the SoxF gene *Sox18* and this perhaps further indicates conserved regulatory functions in epithelial appendage formation (Olsson et al., 2001; Pennisi et al., 2000b). Interestingly the null mutation of *Sox18* in mice only has minor effects on hair formation but the expression of a truncated form produces far more significant phenotypes (Pennisi et al., 2000a). The authors suggest that this is due to a trans-dominant negative effect disrupting the contribution to the regulatory network of multiple *Soxes* and may indicate a more significant role for *Sox2* in this site. This is discussed further in chapter 4.

The diversity of the *Sox* family means that there are very few tissues where there is unique expression of any single member and this may be key to aspects of the regulatory role these genes have to play. There are several examples of tissues where the initial expression of a particular *Sox* gene is augmented or replaced by that of another family member as cells become increasingly committed to a particular fate. Whilst cooperative binding ensures certain targets will be regulated exclusively by one *Sox* gene or another, a subset of targets may be differentially regulated in the presence of various combinations of *Soxes* and their partners. Where the closest relatives of a family have overlapping expression questions regarding shared function arise. Overlapping expression with more diverse relatives might indicate a degree of antagonism or cooperation that is critical for the optimal function of that cell.

Expression of nearly half of all the *Sox* genes has been detected in the developing CNS. The *SoxC* genes *Sox4* and *Sox11* have a more dynamic expression pattern than the *SoxB1* members and tend to be in regions containing differentiating neurons and glia (Cheung et al., 2000). At 16.5dpc in the dorsal telencephalon *Sox11* expression overlaps with that of the *SoxB1* genes and is highest in the sub-ventricular zone where there are late generated neuronal and glial precursor cells. *Sox4* is not found here, but is strongly expressed in dorsal cortical plate that at this stage is made up of early generated cortical neurons. *In vitro* all these genes have the capacity to differentially regulate expression via identical target sequences in combination with a variety of POU domain cofactors (Chapter 1) (Wiebe et al., 2003). It would therefore be interesting to see if the same phenomena occurs *in vivo* and what the consequences are of adjusting the relative amounts of these factors in certain cell types.

Several members of the *Sox* family are also detected in the early stages of inner ear development. The *SoxE* gene *Sox9* is expressed uniformly throughout the otic placode similar to *Sox2,* but by 9.5dpc expression is resolved to the dorsal half the otic vesicle, opposite to that of *Sox2* (Ng et al., 1997). Also at 9.5dpc the *SoxD* gene *Sox6* is expressed by cells of the inner layer of otic epithelium and is then down regulated (Murakami et al., 2001). Another *SoxD* gene *Sox13* is expressed at 13.5dpc in the maculae (utricle and sacculae) but not in other structures (Roose et al., 1998). *In vivo*, different combinations of these (and other) factors in a single cell may result in competition for a subset of common targets and could result in a wide range of expression profiles that may trigger particular fate decisions.

The *Sox* genes employ a number of strategies to promote appropriate target specification and regulation. However, given that some *Sox* genes, even from different subgroups, are capable of differentially regulating common targets, optimal expression of a subset of genes may rely on a particular combination of related transcription factors (Kamachi et al., 1995; Wiebe et al., 2003). The idea that a cell's fate may be dictated by the balance between the activities of *Sox* subgroup members has been suggested previously (Uchikawa et al., 1999). The existence of a more general regulatory strategy where diverse combinations of *Sox* genes and other factors regulate common targets *in vivo* is a point worthy of further investigation. In this way a relatively small number of similar regulatory factors has the potential to create the diversity of expression required for the development of complex organisms.

The diverse pattern of *Sox2* expression may be useful in identifying candidate targets or regulatory partners important for its function during development. In mouse *Sox2* expression is matched to some extent by the mammalian *hairy* and *enhancer-of-split* homologues *Hes1* and *Hes5*. During CNS development these basic helix-loop-helix transcription factors share sites of expression with *Sox2* in the olfactory bulb, cerebellum and the neuroepithelial lining of the ventricles (Akazawa et al., 1992). Similar expression is also observed in the developing sense organs and in the bulge region of vibrissal and monotrick hair follicles (Cau et al., 2000; Hojo et al., 2000; Leonard et al., 2002; Zheng et al., 2000). Targeted disruption of these genes indicates that they inhibit neuronal differentiation by mediating *notch* signalling and mutation is most severe in embryos simultaneously null for both of these genes (Ohtsuka et al., 1999). In the cochlea, null mutation of *Hes1* and *Hes5* results in supernumerary inner and outer hair cells respectively (Zine et al., 2001). *Sox2* has also been implicated in maintaining an undifferentiated state and as such may have some role in the expression or function of the *Hes* genes (Avilion et al., 2003; Graham et al., 2003; Li et al., 1998)..

This study extends the current description of *Sox2* expression and shows that it is present in a variety of tissues during the development of the mouse embryo. Because of this diversity it is difficult to assign a particular role to *Sox2* although it is possible that common regulatory mechanisms exist and so information gleaned in one system may aid understanding in a seemingly unrelated tissue. Examining the role of *Sox2* in the developing mouse CNS by genetic manipulation may be very difficult as integrity of this structure is a requirement for successful breeding and maintainance of mouse lines. The disruption or the complete absence of the inner ear or the hair

follicle, however, is unlikely to lead to lethality. These sites therefore offer discrete model systems where the effects of modifying *Sox* gene expression *in vivo* may be relatively easily assessed.

Chapter 3 Conditional mutation of *Sox2* in ES cells.

3.2 Introduction.

The examination and comparison of gene expression patterns is critical in order to formulate hypotheses about which cells and systems may require the action of a particular gene product. Experimental strategies can then be created in order to test these theories and provide evidence of the actual function of a particular gene. One of the most commonly employed methods is to examine the consequences of modifying the amount of a particular gene product within a cell, tissue or even whole animal.

Transcription factors play key roles in the regulatory networks that dictate the overall profile of gene expression in a particular cell type. If the amount of a transcription factor can either be raised or lowered then clues to its place in the regulatory network may be recovered. A favourite approach to achieve this in mice is to study the effect of removing a particular gene by the null mutation of one or both of its alleles.

The study of gene function by null mutation has become a standard method in many organisms. In mouse, null mutants can be created in a random fashion or in a defined, targeted way. Random loss of function mutants may be produced when the gene of interest is unknown. Speculative mutation of the whole mouse genome by chemical or radioactive means produces mouse models that are selected on the basis of some measurable phenotype in a tissue of interest. This is followed by identification of the mutation responsible and which gene or genes are affected. Targeted modification of DNA sequence in the mouse adopts the reverse strategy

whereby a known gene is modified and then there is a search for phenotypes that result from this modification. This method is far more refined and relies heavily on the careful design and production of a suitable targeting vector to facilitate the introduction of a specific genetic mutation and is followed by detailed phenotypic analysis. In essence the targeting vector is a mutated version of a particular allele and is produced by manipulating DNA sequences *in vitro*. This modified version of the allele of interest can be used to specifically alter its genomic counterpart in mice (Smithies and Powers, 1986; Thomas and Capecchi, 1987) (reviewed by (Muller, 1999). This is done by taking advantage of the pluripotent characteristics of embryonic stem cells (ES cells). ES cells are derived from the inner cell mass of blastocyst stage mouse embryos and can be readily manipulated in tissue culture, but also have the ability to contribute to all tissues of a developing embryo, producing chimeric mice, when injected into the cavity of a host blastocyst (Evans and Kaufman, 1981). In this way a mutation introduced in tissue culture can be passed though the germline of chimeric mice to produce lines of animals that carry the engineered modification.

In order to work successfully the targeting vector must contain certain critical components. The desired mutation should be flanked by DNA sequence homologous to the allele to be modified so that recombination with the target region can occur. The modification itself must contain a selectable marker so that, once produced, the few ES cells carrying the mutant allele may be separated from the majority that do not. Once constructed this version of the allele can then be transfected into ES cells in culture and will occasionally replace the normal allele by the process of homologous recombination. After selection and confirmation that the DNA sequence

of a particular ES cell clone has been altered correctly, ES cells can be taken and used to produce mouse lines by injecting them into host blastocysts. Colonization of the host by ES cells will then, hopefully, create a mouse in which manipulated cells are represented in each tissue and in particular contribute toward the germ line. If the modification is able to pass through the germ line then lines of mice may be created that carry the altered allele in each cell. In this manner a specific gene can be modified or removed followed by the examination of any resulting phenotype. The modification most often adopted is the complete or partial removal of the coding sequence so that no functional gene product can ever be produced from that allele.

By using the original replacement method a targeted modification is carried by every cell from the earliest stages of development. Animals harbouring lethal mutations can only be studied up until the point at which the specific alteration is lethal and the earlier this lethality is the fewer the sites are that can be studied. The null mutation of *Sox2* (*Sox2*$^{\beta geo}$) was produced in ES cells by replacing 3.5kb of DNA, including the entire coding region of *Sox2*, with the reporter/selection cassette *βgeo*. These cells were then introduced into blastocysts to produce mouse lines (Avilion et al., 2003). Since mice carrying a homozygous null mutation of *Sox2* die at the peri-implantation stage due to a failure of the epiblast this model can not provide information about the role of *Sox2* in other sites later in development (Avilion et al., 2003).

Sox2 is expressed in a variety of tissues throughout embryonic development as well as in the adult mouse and so a mutation strategy avoiding early lethality would enable an examination of its contribution at later stages (Chapter 2). Ideally

such a strategy allows the removal of *Sox2* in a spatially and temporally restricted manner leaving other sites with normal levels of expression. This can be achieved by using the *cre/loxP* site-specific recombination system (Gu et al., 1993). This method relies upon the ability of the P1 bacteriophage protein *cre* to catalyse recombination between two 34bp *loxP* (locus of 'x' over in P1) sequences placed anywhere in the genome. If a targeted modification produces a region flanked by *loxP* sites, arranged in the same orientation (floxed), intervening DNA may be looped out and consequently removed upon the introduction of *cre* recombinase (Fig. 3.1). Targeting vectors are therefore designed in order to produce functional but floxed regions of DNA in mice generated from ES cells. Conditional removal of the floxed sequence can subsequently be achieved by introducing *cre* either by transfection or breeding (Reviewed by Lewandoski, 2001).

If the floxed sequence contains all or part of the coding sequence of a gene then a null mutation may be produced upon the introduction of *cre*. This technology therefore provides an enormous advance in precision and flexibility when studying gene function by null mutation. There is a growing resource of mouse lines available that express the *cre* protein in a variety of tissues at numerous developmental and adult stages (http://www.mshri.on.ca/nagy/Cre-pub.html). By breeding *cre* expressing mouse lines against those carrying particular floxed alleles specific null mutations can be produced at any stage in virtually any tissue.

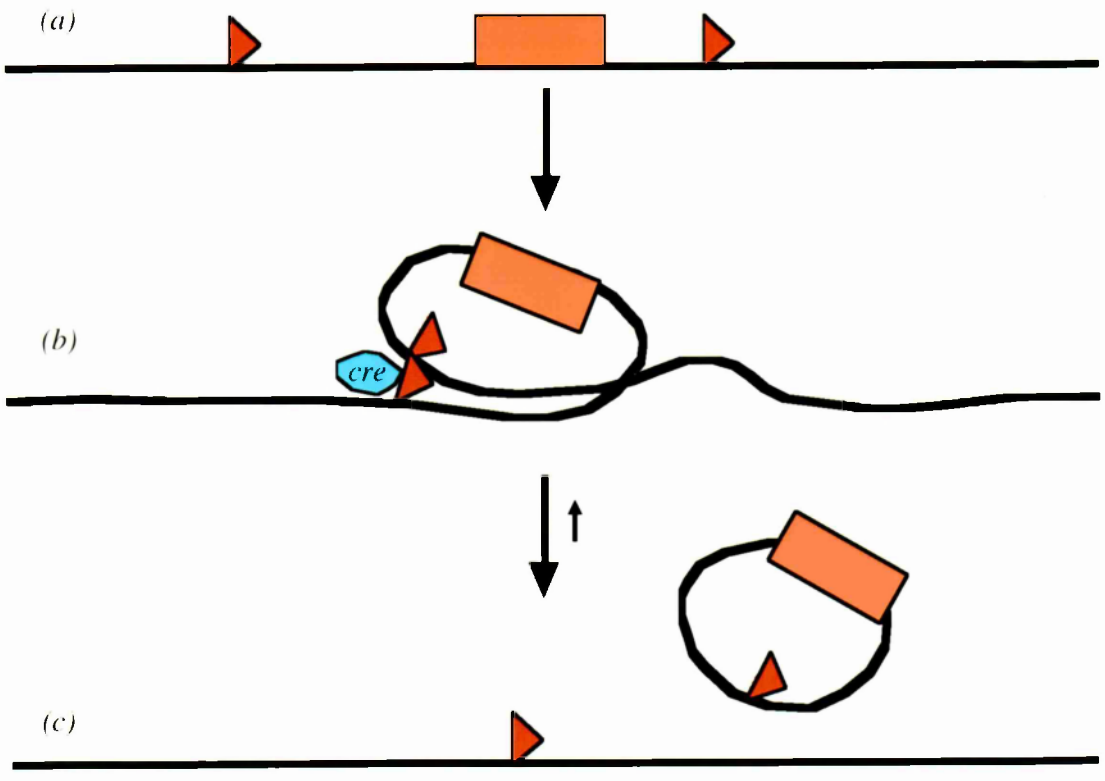
The aim of this experiment was to examine the role of *Sox2*, by null mutation, in some of its diverse sites of expression during the development of the mouse embryo. In order to facilitate this a strategy was devised in order to introduce *loxP*

Figure 3.1 *cre/loxP* site specific recombination system.

(a). *An allele is produced so that a region of interest is flanked by loxP cre recognition sites (floxed).*

(b). *The presence of cre recombinase catalyses association and recombination between the two loxP sequences*

(c). *The two loxP sites recombine, looping out the intervening DNA sequence, which will eventually degrade or be deleted as cells divide. The allele of interest is left without the intervening DNA sequence and a single loxP at the site of recombination. This reaction may also occur in the reverse direction but as the two loxP sites are no longer close to each other in cis this is likely to be a rare event.*


(a)
(b)
cre
(c)

sites into the *Sox2* allele so that its coding sequence can be removed upon the introduction of *cre* recombinase.

3.3 Results.

3.3.1 Targeting strategy using pDB13

The initial strategy employed to produce a floxed *Sox2* allele utilised the targeting vector pDB13 that carries a promotorless selection cassette just upstream of the initiator codon of the *Sox2* coding sequence. *Sox2* is expressed in ES cells and by hijacking the natural *Sox2* promoter this vector would confer resistance to the antibiotic G418 only when correct targeting of one of the *Sox2* alleles has occurred or in the rare event of random integration next to another promoter also expressed in ES cells. This strategy is in essence the same as that used to produce the $Sox2^{\beta geo}$ null allele, which was very efficient as it increased the proportion of correctly targeted ES cell clones after antibiotic selection (Fig. 3.2d) (Chapter 1), (Avilion et al., 2003). The selection cassette itself was flanked by *FRT* (Flp recognition target) sites (flrted) that are the recognition sequences of the *Saccharomyces cerevisiae* site-specific recombinase Flp. The Flp/*FRT* system works in a way that is analogous to *cre*/*loxP* and was employed so that the selection cassette can be removed once correct targeting had been achieved, either by introducing Flp recombinase into cells *in vitro* or by the pronuclear injection of zygotes with a *FRT* expressing construct (Fig. 3.2d,e). Removal of the selection cassette would then leave a floxed *Sox2* gene that is fully functional. The addition of *cre* recombinase will remove the loxP-flanked *Sox2* gene (Fig. 3.2e,f). The *loxP* sites that flanked the *Sox2* coding sequences were placed within the untranslated region (UTR) of the gene in an attempt to minimise the risk of introducing an unintentional mutation that might affect the regulation of *Sox2* (Fig. 3.5).

Figure 3.1 Conditional targeting strategy using pDB13.

Figure summarising the key features of the targeting strategy using pDB13. A more detailed description of the production of this vector is given in materials and methods.

(a). *The targeting vector used to produce the Sox2$^{\beta geo}$ allele. The selection cassette ßgeo replaces approximately 3.5kb of the Sox2 gene including the entire coding region.*

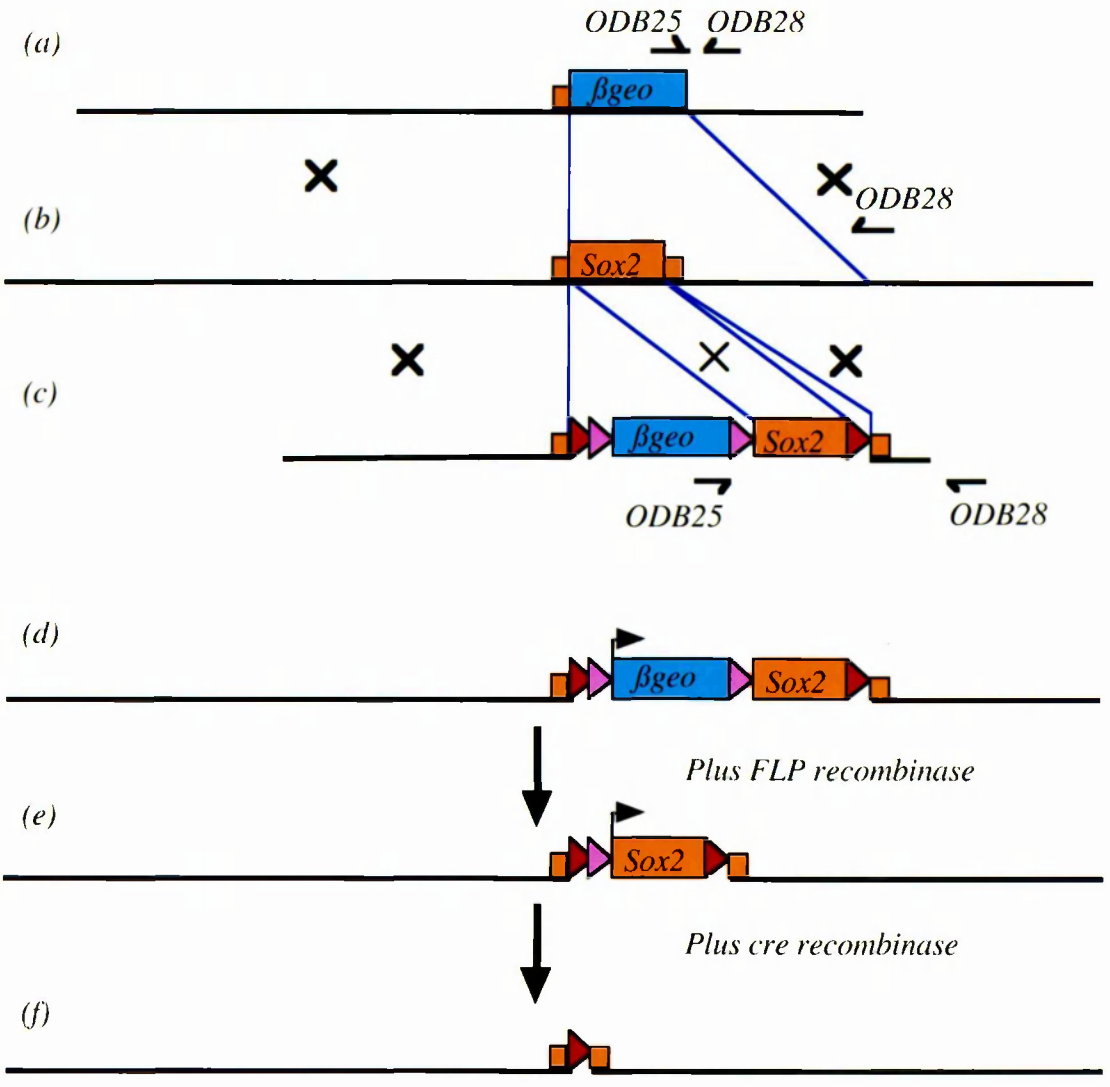
(b). *The Sox2 locus showing regions of homology, where recombination may occur, with both the Sox2$^{\beta geo}$ vector and pDB13.*

(c). *The targeting vector pDB13. The selection cassette is placed so that it inserts at the same point as in the Sox2$^{\beta geo}$ allele but the Sox2 coding sequence and downstream DNA is left intact. The selection cassette is flrted to facilitate its removal once targeting has been achieved.*

(d). *A representation of the Sox2^{flox} allele once homologous recombination has taken place.*

(e). *After the introduction of Flp recombinase recombination between the FRT sites results in the removal of the ßgeo selection cassette. A functional but floxed Sox2 allele remains along with a single FRT site.*

(f). *Upon the addition of cre, recombination occurs between the loxP sites removing the Sox2 coding sequence.*



Region of homologous recombination

loxP, cre recombinase recognition site

Frt, FLP recombinase recognition site

Sox2 — Sox2 coding region

βgeo — Selection cassette

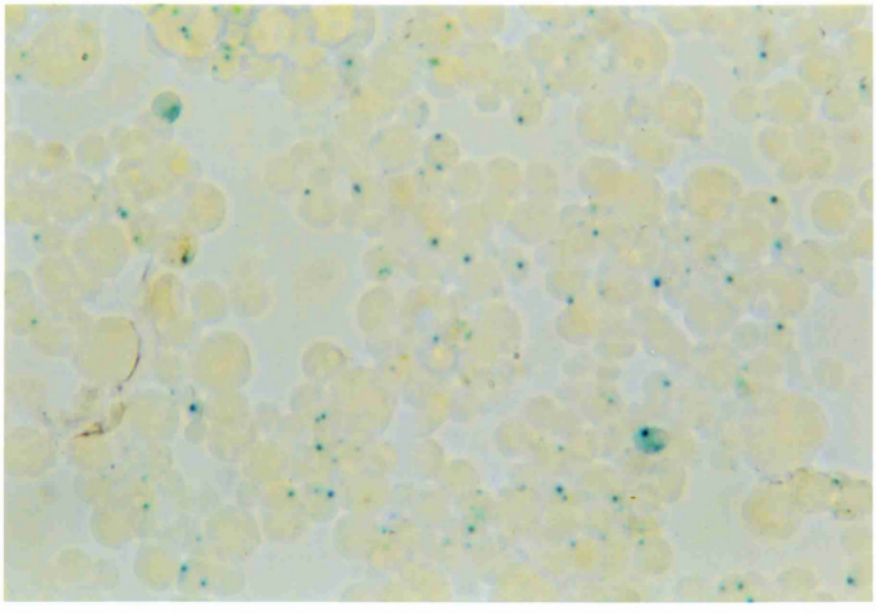
PCR primer

3.3.2 Transfection and screening of ES cells with pDB13

The targeting construct pDB13 (Fig. 3.2c) was made as described in materials and methods and used to transfect CCE ES cells, followed by G418 antibiotic selection. Initially two transfections ($4x10^7$ cells) were performed (materials and methods) and from these there were fifteen colonies of a suitable size to be expanded further for screening. This number was low even for a promoterless targeting vector. Only one of the colonies showed the expected morphology of a correctly selected clone and this was also the only one positive for the targeting vector as judged by PCR (materials and methods). Southern analysis of DNA from this colony reveals that cells were indeed correctly targeted but the modified *Sox2* allele did not contain the second (3') *loxP* site (data not shown). When these cells were stained for *βgeo* activity it revealed a spotty staining pattern where each ES cell showed a very localized spot of colour that seemed to be located at the edge of the nucleus (Fig. 3.3). This indicates that the selection cassette was not performing effectively for some reason but was still capable of producing a protein that could convey some resistance to G418 and catalyse the reaction with X-gal. A further two transfections were carried out where the concentration of antibiotic used in the selection step was reduced to 100μg/ml in an attempt to encourage the expansion of targeted clones expressing an inefficient version of the selectable marker. This strategy however produced no further suitable ES cell clones and was therefore abandoned in favour of constructing a new vector.

Figure 3.3 X-gal staining of the single clone isolated from pDB13 targeting.

pDB13 clone 1



3.3.3 Targeting strategy using pDB37.

A second targeting vector (pDB37) was designed using an alternative strategy to facilitate the conditional disruption of *Sox2*. The second strategy is more conservative than the first in order to increase the chance of success but it required the incorporation of certain components that could potentially interfere with the aim of the experiment. The new strategy abandoned the promoterless selection cassette in favour of using a neomycin gene (*neo*) expressed under the control of the ubiquitously expressing phosphoglycerate kinase (PGK) promoter (PGK*neo*). The function of this selection cassette had already been tested in other targeting experiments but there are potential problems associated with its promoter. The PGK promoter is able to influence the regulation of genes in the vicinity of its insertion site and so unexpected phenotypes my arise from its use (reviewed by Muller, 1999). To minimise any potential interference during the targeting and selection steps the PGK driven selectable marker was placed in the opposite orientation, downstream of the coding region of interest. Subsequent interference may be avoided as the entire cassette is flrted and so can be removed when no longer required (Fig. 3.4d,e). Upon the introduction of Flp recombinase only a single *loxP* and an *FRT* site remain and are not likely to cause any significant problems (Fig. 3.4e).

Suitable insertion sites for the selection cassette and distant *loxP* sequence were carefully identified so as to minimise the possibility of interference with the normal function of the modified allele. A search was carried out for regions of homology between mouse and the human genomic sequence available at the time in the vicinity of the *Sox2* coding sequence. Regions that were highly conserved

Figure 3.3 Conditional targeting strategy using pDB37.

Figure summarising the key features of the targeting strategy using pDB37. A more detailed description of the production of this vector is given in materials and methods.

(a). *The targeting vector used to produce the* $Sox2^{\beta geo}$ *allele. The selection cassette ßgeo replaces approximately 3.5kb of the Sox2 gene including the entire coding region.*

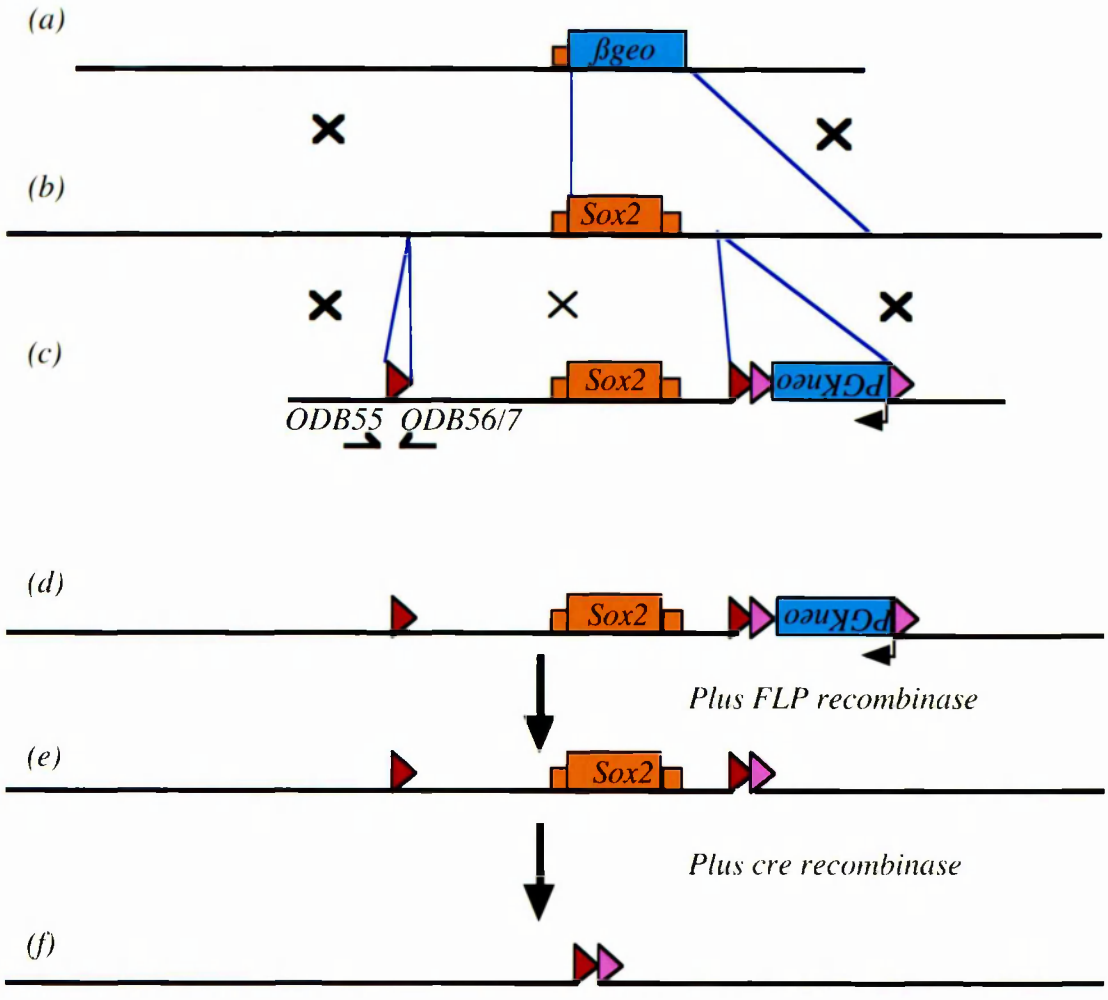
(b). *The Sox2 locus showing regions of homology, where recombination may occur, with both the* $Sox2^{\beta geo}$ *vector and pDB37.*

(c). *The targeting vector pDB37. The flrted selection cassette was placed downstream and in the opposite orientation to Sox2. The second loxP site was placed some distance upstream, avoiding possible regulatory sequences.*

(d). *The initial Sox2 allele that should result from a correct targeting using pDB37.*

(e). *Upon the introduction of Flp recombinase the selection cassette should be removed leaving a single FRT and a single* loxP *site downstream of the using coding sequence.*

(f). *When* cre *recombinase is introduced the entire coding region of Sox2 will be removed.*



Region of homologous recombination

loxP, cre recombinase recognition site

Frt, FLP recombinase recognition site

Sox2 coding region

βgeo or PGKneo Selection cassette

PCR primer

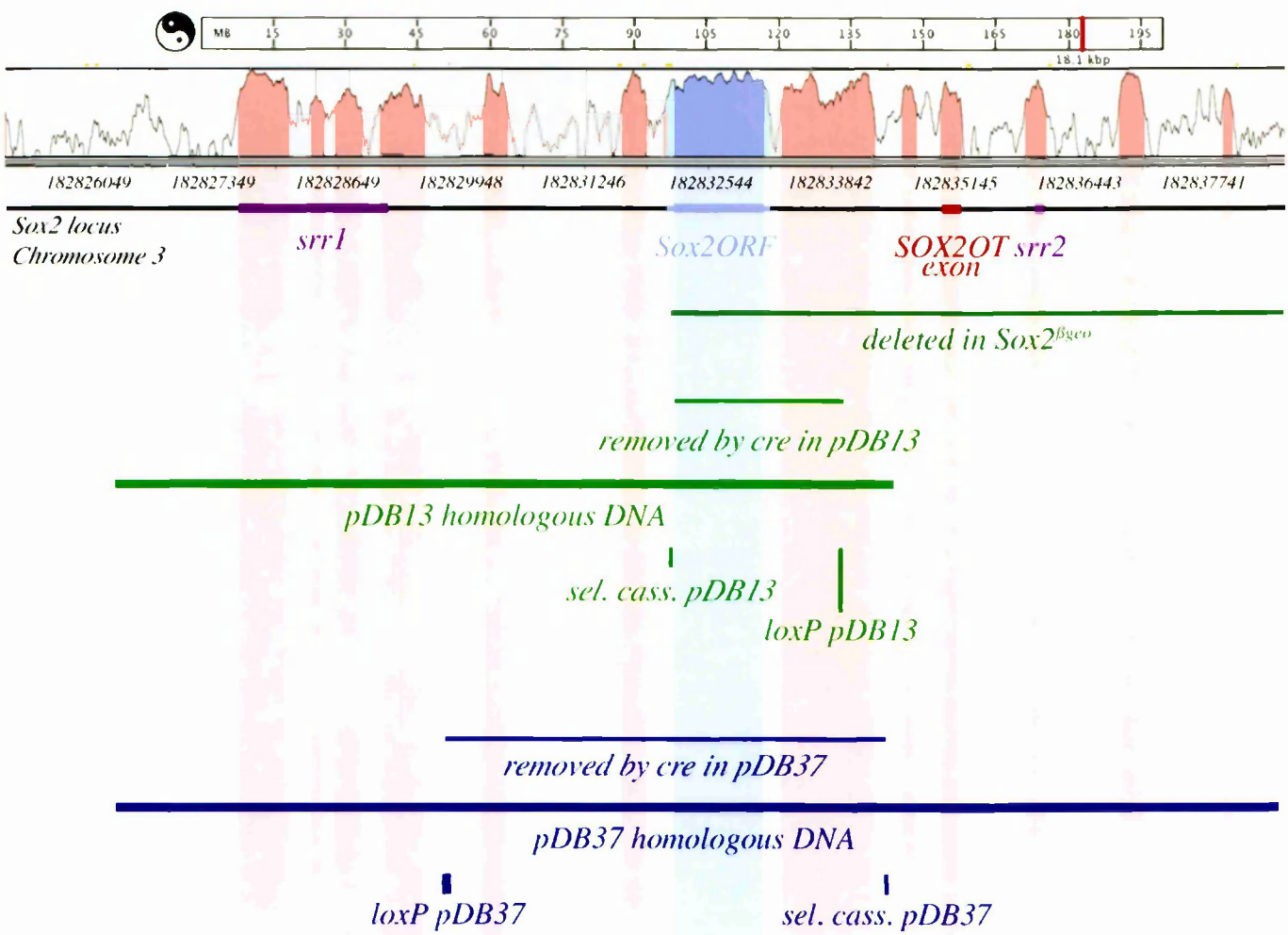
between the species might have some function and so were avoided. Other regions that showed DNase hypersensitivity and so might be involved in gene regulation were also avoided (Zappone et al., 2000). The selection cassette (and first *loxP*) was placed just downstream (50bp) of the transcribed region of *Sox2* in a stretch of relatively unconserved sequence (Fig. 3.5). Using the same criteria the most suitable site for the flanking *loxP* was considered to be about 4.5kb away from the selection cassette upstream of the initiating codon of *Sox2*.

Regions of DNase hypersensitivity have subsequently been refined to more specific sequences that are involved in the regulation of *Sox2* expression (Tomioka et al., 2002). The more complete genomic sequence available at the time of writing shows extensive identity between human, mouse and chicken sequence both upstream and downstream of the *Sox2* open reading frame (Uchikawa et al., 2003). The insertion sites for the selection cassette and the second *loxP* site, however, remain in relatively unconserved regions (Fig. 3.5).

The choices made during the design of pDB37 are likely to produce a fully functional, floxed *Sox2* allele but at the expense of a low efficiency targeting and selection step. Firstly far more antibiotic resistant clones are produced by using a ubiquitously expressed selectable marker with most being random integrants and therefore not useful. Secondly the distance between the two *loxP* sites increases the chance that when targeting does occur, homologous recombination will exclude the second *loxP* as there is no selection pressure favouring its presence. In an attempt to increase targeting efficiency a longer stretch of homologous sequence was used than

Figure 3.4 Summary of *Sox2* allele with insertion sites and published hotspots.

A diagram relating the targeting constructs pDB13 and PDB37 to the Sox2 locus. Vista alignment (Visualization Tools for Alignment, http://www-gsd.lbl.gov/vista/VISTAdownload2.html) highlights identity between mouse and human genomic sequence once repetitive sequence has been removed. Top bar adjacent to ying-yang represents human chromosome three with red bar indicating position of comparison. Identity greater than 50% over a 100bp window is shown with identity greater than 75% highlighted in pink. Regions where sequence is significantly conserved may have an important function and are best avoided when targeting the locus. The deletion caused by the Sox2$^{\beta geo}$ vector removes the Sox2 coding sequence but also one of the SOX2OT exons. Regulatory sequences srr1 and srr2, described by Tomioka et al, are also shown and indicate just two of regulatory sites best avoided in a conditional strategy (Tomioka et al., 2002).


MB 15 30 45 60 75 90 105 120 135 150 165 180 195
18.1 kbp
182826049 182827349 182828649 182829948 182831246 182832544 182833842 182835145 182836443 182837741
Sox2 locus
Chromosome 3
srr1
Sox2ORF
SOX2OT srr2
exon
deleted in Sox2βgeo
removed by cre in pDB13
pDB13 homologous DNA
sel. cass. pDB13
loxP pDB13
removed by cre in pDB37
pDB37 homologous DNA
loxP pDB37
sel. cass. pDB37

with pDB13 (Fig. 3.5), even so many colonies still needed to be screened in order to isolate those with the complete modification.

The targeting vector pDB37 was transfected into AB1 ES cells (kind gift R Behringer) and after exposure to G418 antibiotic, twelve hundred colonies were selected for expansion and screening. Out of the twelve hundred, fourteen colonies were deemed as correctly targeted by Southern blot, using a 3' probe, external to the targeting vector (Fig. 3.6a,b). Out of those fourteen only two had the desired recombination events with the second *loxP* site also being present (Fig. 3.6c). These two clones (8B9 and 10D6) were further characterised before being used for the production of chimeras.

The clones selected for injection grew normally in culture and in total twenty-six chimeras were produced from their injection into C57Bl/6 host blastocysts. Generally there was a good contribution from both ES cell clones with eight (30%) of the chimeras produced having greater than ninety percent ES cell contribution as judged by agouti coat colour (Table 3.1). There was also a sex ratio distortion in mice derived from these cells with twenty-two of the chimeras produced being male. This gave a percentage ratio of 85:15 male to female rather than equal numbers of each.

Sex ratio distortion and ES cell quality.

The sex ratio distortion suggests that some of the animals derived from XX host blastocysts had been converted to males due to the contribution of the XY ES cells injected into them. If the proportion of XY ES cell derived cells in the developing gonad is high enough then the chimera will be phenotypically male,

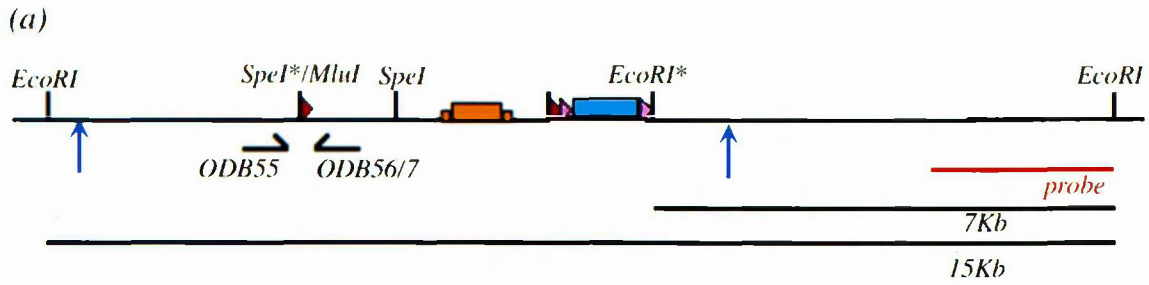
Figure 3.5 Screening of ES cell clones derived from transfection with pDB37.

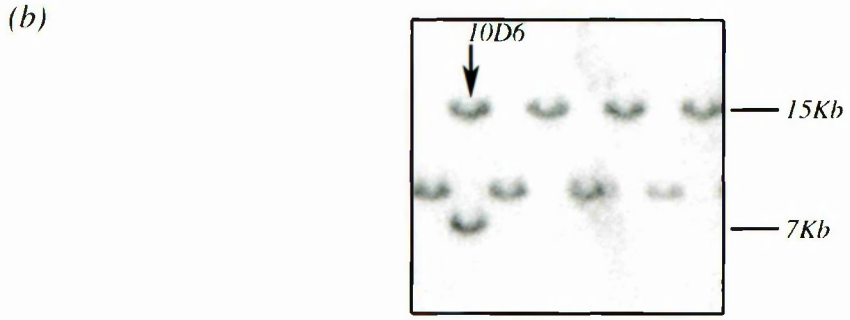
(a). *Diagram representing a Sox2 allele correctly targeted after homologous recombination with pDB37. Restriction enzyme sites shown are those useful for screening purposes with '*'indicating those sites that would be introduced by modification of the endogenous Sox2 allele. Blue arrows indicate the ends of homologous sequence in the pDB37 construct. Southern blots were performed by digesting genomic DNA with EcoRI. The fragmant indicated in red is not contained within the targeting construct and could redily detect wildtype or targeted Sox2 alleles which were 15kb or 7kb respectively. PCR screening of positive samples with the oligonucleotide pairs ODB55/56 and ODB55/57 was used to ensure that loxP site distant from the selection cassette had also been encorporated.*

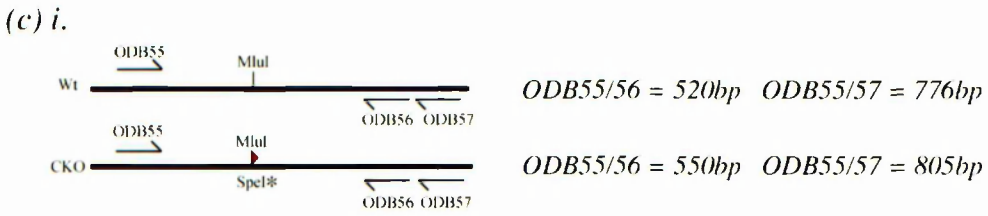
(b). *ES cell genomic DNA digested with EcoRI and probed with a fragment external to the targeting construct pDB37 can clearly distinguish between wildtype and targeted allele. Here eight staggered lanes are shown. Only the lane with 10D6 DNA has 15kb and 7kb bands characteristic of a wildtype and a targeted allele respectively.*

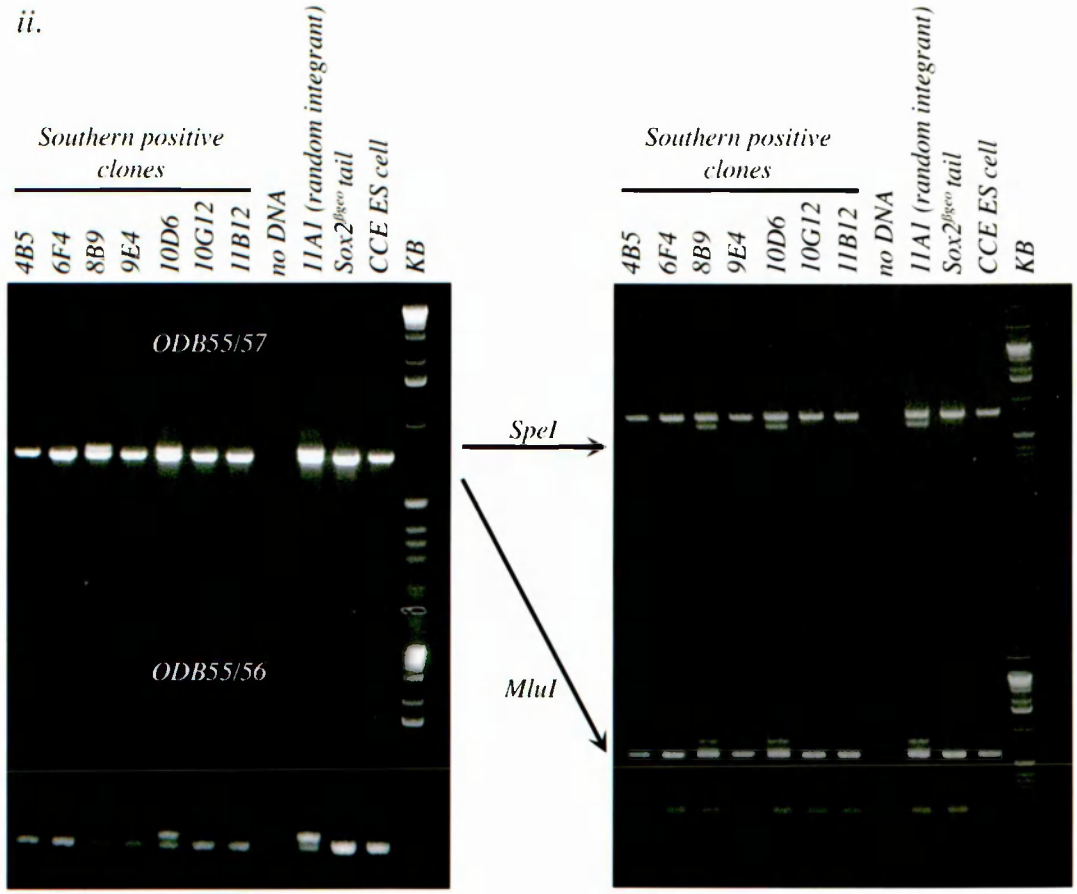
(c)i. *PCR was utilised to screen targeted ES cell clones for the loxP site distant from the selection cassette. Two alternative pairs of oligonucleotides could each be used to discern between wildtype (wt) and targeted alleles.*

(c)ii. *ES cell clones deemed correctly targeted by Southern blotting using a 3' probe were screened with PCR to establish if homologous recombination had also included the distant loxP site. Two of these clones showed characteristic double bands indicating complete targeting. Enzymatic digestion of the PCR products supports the existance of the intended mutation. Marker DNA is 1kb ladder from Boehringer Mannheim.*


(a)
EcoRI
SpeI*/MluI
SpeI
EcoRI*
EcoRI
ODB55
ODB56/7
probe
7Kb
15Kb


(b)
10D6
15Kb
7Kb


(c) i.
ODB55
MluI
Wt
ODB56
ODB57
ODB55
MluI
CKO
SpeI*
ODB56
ODB57
ODB55/56 = 520bp ODB55/57 = 776bp
ODB55/56 = 550bp ODB55/57 = 805bp


ii.
Southern positive clones
4B5
6F4
8B9
9E4
10D6
10G12
11B12
no DNA
11A1 (random integrant)
Sox2flgeo tail
CCE ES cell
KB
ODB55/57
ODB55/56
SpeI
MluI
Southern positive clones
4B5
6F4
8B9
9E4
10D6
10G12
11B12
no DNA
11A1 (random integrant)
Sox2flgeo tail
CCE ES cell
KB

although it is actually a mixture of XX and XY cells. In both phenotypically and genotypically female or genotypically male chimeric gonads, gametes may arise from cells that are either derived from the host blastocyst or the XY ES cells, giving black or agouti pups respectively. In a chimera which is phenotypically male, due to a large XY ES cell contribution, but derived from a genotypically female host blastocyst, wild type XX cells are not able to give rise to viable male gametes and so any functional sperm in these animals must be derived from the initial ES cell contribution. The fertility of these mice will therefore depend entirely on the quality of the ES cells used to produce them and so will be either infertile or give rise only to agouti pups.

Breeding of chimeras.

The chimeric mice, made from the two independently targeted ES cell clones, were bred against MF1 animals (albino). Coat colour was examined in the offspring of chimeras, with agouti pups indicating a transmission of ES cell derived genetic material. Out of 974 pups produced by the chimeras none had agouti coat colour and litter sizes were normal (Table 3.1). Also five male chimeras, (which may represent sex converted ones), were not able to make MF1 dams pregnant despite evidence of normal mating. These results suggest that the ES cell clones used to generate the chimeras are unable to contribute to the germ cells.

Table 3.1 Chimeric mice and offspring created.

Chimeric mice produced by the injection of targeted ES cells into host blastocysts were selected on the basis of coat colour. The proportion of agouti fur indicates the degree of contribution of ES cell derivatives in these animals. Chimeric mice were then bred against wild-type animals with the hope that the targeted allele would pass through the germline. In the event of germline transmission offspring should have a completely agouti coat.

Line	ID number	%Chimerism	Sex	Total Offspring	Total Agouti
8B9	11	90	Male	0	0
8B9	14	85	Male	0	0
8B9	19	65	Male	41	0
8B9	13	60	Male	34	0
8B9	18	60	Male	39	0
8B9	15	45	Male	54	0
8B9	12	30	Male	66	0
8B9	16	20	Male	38	0
8B9	17	20	Male	30	0
8B9	20	20	Male	50	0
8B9	23	20	Female	9	0
8B9	22	15	Female	16	0
8B9	21	5	Female	9	0
10D6	7	95	Male	37	0
10D6	8	95	Male	82	0
10D6	24	95	Male	27	0
10D6	25	95	Male	27	0
10D6	26	95	Male	0	0
10D6	2	90	Male	0	0
10D6	5	90	Male	35	0
10D6	9	70	Male	91	0
10D6	1	30	Male	101	0
10D6	4	15	Male	113	0
10D6	6	10	Male	75	0
10D6	3	5	Female	0	0
10D6	27	5	Male	0	0
				974	0

3.4 Discussion

The expression study (Chapter 2) indicates that *Sox2* plays a significant role in a number of tissues of the developing mouse embryo after the point at which the existing homozygous null mutants die. Production of a conditional null mutation would therefore offer a valuable resource with which to study the role of *Sox2* throughout murine developmental and adult stages. The results of this chapter demonstrate that although gene targeting techniques promise a powerful method by which gene function can be studied these approaches can be fraught with technical difficulties. Ultimately, however, it is worth pursuing such strategies since they offer precise ways of manipulating the mouse genome in order to create *in vivo* models.

In the initial strategy pDB13 was designed to maximise targeting efficiency whilst minimising the chance of the required modifications causing aberrant phenotypes. The promotorless strategy used here is very similar to that employed to produce the $Sox2^{\beta geo}$ allele, which had a very good targeting efficiency. In the $Sox2^{\beta geo}$ targeting experiment forty G418 resistant ES cell colonies were selected for expansion and screening yielding seven correctly targeted clones so it was peculiar that the pDB13 strategy was so unsuccessful. pDB13 was able to modify the *Sox2* allele since the only colony to survive was indeed targeted correctly although there was an unusual *βgeo* staining pattern. The extremely low targeting efficiency and lack of G418 resistant colonies arising due to random integration suggests that there is a problem with the activity of the selection cassette. There could be a number of possible explanations for this including mutation, misregulation or a disruption of proper transcription or translation of the selectable marker. Observations suggest

that both components of the selection cassette do indeed work, at least to some extent, since one colony was blue and G418 resistant. The site of the selection cassette within the targeting vector should permit *ßgeo* expression in ES cells as it did not differ from the original $Sox2^{\beta geo}$ targeting construct and so there is not likely to be a problem with the promoterless strategy. The location of the second *loxP* site is also unlikely be problematic since it lies within the 3'UTR of *Sox2* and therefore is probably separate from any *Sox2* regulatory elements required for the expression of the selectable marker in ES cells. Additionally the exclusion of this region of the UTR in the previous $Sox2^{\beta geo}$ targeting experiment had no adverse affect on expression of the selectable marker in ES cell culture.

The most notable difference between the conditional targeting construct and the original is that pDB13 contains a pair of recombinase recognition sites just before the start codon of the *ßgeo* cassette. The sequence of the recombinase recognition sites include repeats of core recognition sequences. Although these have been placed individually in the transcribed region of other genes, without affecting correct expression, placing both the *loxP* and *FRT* sequences next to each other may have the effect of interfering with efficient transcription or translation of the selection cassette. Anecdotally this idea is supported by the lack of random integrants produced in the four transfections attempted with this construct when they should represent a majority. Even dramatically reducing the concentration of G418 to encourage the growth of clones with a low level of antibiotic resistance had no effect.

Although both the *ßgal* and the *neo* components of the selection cassette had some function and homologous regions were able to direct the modified gene to one

of the *Sox2* alleles it was not practicable to pursue the pDB13 strategy. The experimental aim was to investigate the role of *Sox2* using a conditional mutation, it was therefore most efficient to design and construct an entirely new targeting vector.

The second strategy, using pDB37, was successful and was designed primarily to create a fully functional floxed gene whilst compromising a little on targeting efficiency. The ubiquitous expression of the selectable marker meant that a far greater number of random integrants were produced than would have been with a successful promoterless strategy. Although many ES cell colonies had to be screened the targeting did yield fourteen clones where homologous recombination had taken place and the *Sox2* gene had been targeted. Great care was taken to avoid possible regulatory elements but this introduced further inefficiency due to the distance between the selection cassette and the second *loxP* sequence. Complete targeting was therefore a rare event with correctly modified clones occurring at an overall rate of about one in six hundred. This is not necessarily a problem but for an ideal experiment both clones needed to colonise the germline of chimeric mice and unfortunately they did not.

The two ES cell clones deemed to be correctly targeted both produced high percentage chimeras and this was an indication that the ES cell derived cells of the adult mice are able to contribute and behave normally in many of the tissues of these animals. Extensive breeding, however, did not yield any live animals that carried the floxed *Sox2* allele. The fact that the chimeras produced no ES cell derived agouti pups is an indication that ES cell derived gametes give rise to non-viable embryos or that ES cells were not contributing to a functional germ cell population.

Chimeric animals can support cells with otherwise lethal mutations because of clonal expansion or non-cell autonomous effects of wild-type cells (Bi et al., 1999). Also ES cells injected into host blastocysts do not contribute to extra embryonic components and so mutations affecting these tissues would not become apparent until the next generation is produced. Once all cells of the developing embryo carry a mutation lethal effects can no longer be compensated. In this instance this could explain the lack of agouti pups from chimeric animals but litter sizes were normal perhaps suggesting a problem with fertility rather than lethality of a proportion of embryos.

The majority of chimeras were fertile because of gamete production from cells derived from the host blastocyst. Any gametes produced from this population would produce black pups. It is possible that offspring carrying the targeted gene died in utero due to a dominant effect of the mutated allele. Since many high percentage chimeras were born and these animals produced litters of normal size it is more likely that failure to transmit the floxed *Sox2* allele is due to the failure of these cells to colonize the germ line. There was a proportion of male chimeras that appeared infertile and may have represented XX/XY animals in which no gametes could be produced either by the wild type XX C57Bl/6 or the XY ES cells. There may have been no viable gamete production at all in these phenotypically male animals as judged by breeding although they might have been able to contribute enough functional cells to the gonad to sex reverse the chimeras. SOX2 and SOX3 have been detected during stages of male meiosis (Shanthi Mahadaviah and James Turner unpublished data) and although a single null *Sox2* allele should not affect this

process, mis-expression, caused by regulatory interference, could potentially confound normal gamete production.

It is possible that these ES cell clones are not able to contribute to the germline for reasons that are entirely unrelated to the targeted allele. ES cells are notoriously vulnerable to slight changes in their preferred culture conditions and usually several independently targeted clones are injected to produce chimeric mice with the expectation that not all will be able to contribute to the germ line. In this instance the low efficiency of complete targeting yielded only two clones suitable for expansion and injection into host blastocysts. Since the two clones selected as properly targeted were derived from different plates after electroporation they must represent independent clones. It is therefore unlikely that they carry the same spontaneous mutation that might have arisen from experimental procedure. All tissue culture was carried out meticulously and cells used to produce chimeras were shown to be negative for mycoplasma, a common tissue culture contaminant that is known to interfere with germline transmission (Robertson, 1989). The number of passages in culture was also kept to a minimum since it is known that increased passage number reduces the ability of these cells to contribute to the germ line (Fedorov et al., 1997) . It is possible that there was some problem with the ES cell stock before targeting *Sox2*, however, cells were shown to transmit to the germline in previous experiments in our lab and elsewhere (Bi et al., 1999; Rizzoti et al., 2004). Problems of this nature could be overcome by further transfections with the same construct and injection of newly targeted clones in order to produce mice.

If there is a problem associated with the targeted allele then either *Sox2* or some other gene may be affected by the engineered mutation. This effect is most likely to be mediated by up-regulation of a gene or genes in the vicinity since animals heterozygous for the $Sox2^{\beta geo}$ null mutation, in which this region of DNA is absent, are viable and able to breed. The $Sox2^{\beta geo}$ mutation not only removes the entire coding sequence of *Sox2* but also a further 2.7kb downstream which includes the second exon of a recently identified gene *Sox2OT* (Fantes et al., 2003).

The discovery of *Sox2OT* emphasises the importance of creating the minimal amount of disruption when engineering the genome. *Sox2OT* does not appear to be translated but transcripts are highly conserved between mouse and human and it is suggested to play a role in gene regulation (Fantes et al., 2003). Inadvertently disrupting the function of this or any another gene is a major risk associated with gene targeting and is why both pDB13 and pDB37 were devised to minimise such potential effects both before and after the introduction of *cre*. It is not known if this gene might have some role in germ cell development but it is possible that its regulation is affected by the presence of the PGK promoter and this is why the targeted ES cells are unable to contribute to the germline.

Correct targeting with the vector pDB37 leaves this entire region intact, inserting DNA at only two points flanking the *Sox2* coding region. If there is an inappropriate regulation of *Sox2* or one of the nearby genes, preventing germline transmission, then this is most likely mediated by the presence of the *PGKneo* cassette. There are several examples where placement of the *PGKneo* cassette in intronic, 5' or 3' flanking or in untranslated regions interferes with the normal

function of genes in its vicinity (Lewandoski, 2001; Muller, 1999). *Sox2* is required to maintain the pluripotency of cells in the early embryo and for the derivation of ES cells but the existence of $Sox2^{\beta geo-/+}$ mice demonstrates that a single null allele is not usually lethal (Avilion et al., 2003). Constitutive expression of *Sox2* in ES cells, however, results in extensive cell death (Mitsui et al., 2003). In this instance targeted ES cell clones appear completely normal *in vitro* and they are able to make a significant contribution to chimeric animals. This is a strong indication that targeted clones are able to maintain their pluripotent behaviour and that, at least during early stages, adequate and not excessive amounts of *Sox2* are expressed from engineered cells.

Potential interference could be overcome by removing the *PGKneo* cassette. pDB37 was constructed incorporating a *PGKneo* cassette flanked by *FRT* sequences so that it can be removed after correct targeting if necessary. Targeted clones can be transfected with Flp recombinase *in vitro* with loss of their antibiotic resistance indicating that the selection cassette has been successfully removed (Fig. 3.4e). These ES cell clones should then be checked by Southern analysis to confirm that they have a copy of *Sox2* flanked by *loxP* sites but lack the selection cassette. Floxed ES cell clones could then be injected into blastocysts to produce chimeras and would hopefully facilitate germline transmission of the targeted allele. This would leave an allele that differs from the wild type by the addition of two *loxP* sites and a single stretch of *FRT* sequence representing a total of less than one hundred bases. The majority of this lies within unconserved sequence that is removed completely in the $Sox2^{\beta geo}$ allele and, unless it separates critical regulatory elements, it should hopefully not cause a problem to the normal function of the locus before the introduction of *cre*.

The decision to inject ES cell clones without first removing the selection cassette was made for two main reasons. Firstly, the chances of germline transmission are considered to be better when cells have had a minimum number of passages in culture. Secondly, some degree of interference from the *neo* cassette might have produced a hypomorphic allele that, in combination with the null and wild type, could have been very informative (Nagy et al., 1998). Since *Sox2* heterozygous null animals are viable it is reasonable to assume that, if the selection cassette has the effect of inhibiting the correct expression of this gene to some degree, mice should still be viable. By removing the selection cassette only once a mouse line had been created, an allelic series could be produced including the floxed gene with interfering selection cassette present, a floxed gene and an allele null for the gene of interest (Fig. 3.4d-f).

Compromises made whilst designing pDB37 means that recombination between the two *loxP* sites deletes 4.5kb of genomic DNA including the entire transcribed region of *Sox2* but avoids any *Sox2OT* exons. Although it is formerly possible that part of this sequence is necessary for non-*Sox2* function this likelihood was considered slight although should remain a consideration when examining any null phenotypes.

Once the difficulty of early lethality is circumvented by the production of a floxed *Sox2* allele tremendous prospects toward the study of the role of this gene in a wide variety of embryonic and adult tissues are opened up. Mutations could involve removal of *Sox2* from a very restricted set of cells or tissues thus focusing upon discrete developmental questions. The targeting of a suitable mutation to the *Sox2*

allele should be fairly straight forward. After introducing cre, lines carrying a floxed *Sox2* gene would need to be bred to the second filial generation (F2) in order to produce animals that completely lack a functional *Sox2* allele in a particular organ or tissue. In order to speed up this process, as well as providing a useful marker for *Sox2* expression, animals carrying a $Sox2^{\beta geo}$ allele as well as the *cre* allele could be used. By breeding $Sox2^{flox/flox}$ animals against $Sox2^{\beta geo-/+}/cre$ (tissue X) $^{tg/tg}$ about half the offspring created would be useful, as they would lack a functional *Sox2* allele in tissue X and would express *βgeo* in cells that otherwise should be expressing *Sox2*.

Analysis shows that *Sox2* is expressed in a range of tissues throughout mouse development. These sites can be broadly divided into two categories, those with overlapping expression of *Sox2's* closest family members *Sox1* and *Sox3* and those in which its expression is unique amongst the group B genes. It is possible that *Sox2* contributes similarly towards the regulatory network in all the sites in which it is expressed, although the outcome of this activity may be distinctly different. The disruption of *Sox2* in sites where there is no overlapping expression with either of the other *SoxB1* genes may reveal its role in the development of these tissues but also facilitate the understanding of more general molecular mechanisms that would be difficult to reveal in cells where there may be functional compensation by other genes. There are a number of tissues where the expression of *Sox2* appears to be unique with respect to its closest relatives, for instance in the early sensory placodes, dermal papilla, endoderm or in preimplantation development. *Cre* lines allowing the removal of *Sox2* in these specific sites could be produced.

A unique role of *Sox2* may also be investigated in the earliest populations of cells of the developing embryo. The mechanisms by which embryonic progenitor cells maintain their pluripotency is poorly understood and it has not been possible to fully investigate these mechanisms using the original *Sox2* null mutant. The ZP3-cre mouse (Lewandoski et al., 1997), which expresses *cre* in the developing oocyte, could be used to produce pre-implantation embryos that have no maternal contribution of *Sox2*. This experiment would unmask any previously unobserved early function of SOX2 *in vivo* that is extremely difficult to achieve with traditional null mutations. It will also be possible to produce ES cell lines that are homozygous for the floxed allele and thence study *Sox2* function in cells from the earliest stages of development *in vitro*.

The most significant site of *Sox2* expression is in the developing CNS. There is a considerable overlap of expression in this tissue with that of many members of the *Sox* family, and in particular with those of *group B*, posing the question of whether there is any functional overlap between these genes. It is possible that subtle differences between these genes means that the activity of each one compared to the others in any given cell type might elicit a unique response. Consequently appropriate cell function may rely upon the presence of the correct proportion of these proteins. The perturbation of these proportions could lead to a modification of cell fate. Rather than disrupting multiple genes throughout a tissue as significant and as critical as the CNS, mutation may be limited to a particular region for example part of the hindbrain. For example the *Krox20-cre* mouse (http://www.mshri.on.ca/nagy/Cre-pub.html) could be used to remove *Sox2* from rhombomeres 3 and 5. In this manner the phenotypic consequences would be

restricted allowing groups of mutant cells to be studied in the context of a normally developing embryo.

Another question to address is whether *Sox2* performs slightly different functions in the regulatory networks that exist in early CNS progenitor cells and cells of the more developed embryonic or adult CNS where it is also expressed. Temporally specific disruption of *Sox2* offers the possibility of examining this over several stages of development. For instance *Sox2* is expressed in pluripotent progenitors of the CNS at the earliest stages of nervous system development but also in a similar population of cells lining the ventricles of the brain in late gestation embryos. Although its presence might be a requirement for these cells at each of these stages *Sox2* is not necessarily performing exactly the same function. The disruption of *Sox2* at these different stages using a tissue specific inducible *cre* may therefore expose distinct regulatory functions. This will give insights into how *Sox2* affects the proliferation, migration and differentiation of neural cells in the developing nervous system and the developed brain.

Many possibilities are offered by a floxed *Sox2* allele that justify the laborious nature of its production. In particular this type of genomic manipulation allows subtle modifications to be achieved *in vivo* facilitating the design and execution of increasingly refined experiments.

Chapter 4 Mouse mutants *Lcc* and *Ysb* help to uncover functions of *Sox2* at late stages in embryonic development.

4.2 Introduction

The conditional null mutation of *Sox2* would provide an ideal method by which to study, in a highly defined way, its role in the regulatory networks of particular tissues in the developing mouse embryo. However, other methods of answering questions about the role of *Sox2* in murine development are also available. Whilst creating a conditional null mutant, these methods can be employed to provide further information from a different perspective hence increasing the overall understanding of this gene's function.

Many mouse mutants have been created either by deliberate mutagenesis induced by chemicals or radiation, or due to the random integration of transgenes, often as the serendipitous by-product of other experiments. By their nature these random mutations are difficult to characterise but an increasing list of mutant phenotypes is appearing in the literature (eg SHIRPA http://www.mgu.har.mrc.ac.uk/mut.html). Diverse phenotypes can arise from mutations affecting different genes but some also affect only distinct aspects of a single gene's function. These may range from those that simply remove a gene by rearrangement, to single base pair mutations that modify function due to codon change. Codon changes can result in premature truncation of a protein or perhaps alter its activity. The action of these methods of mutagenesis is not however restricted to those parts of a gene that encode the functional protein. Regulatory elements can be disrupted producing, mice that lack the expression of a particular

gene, but only in a limited range of its expression sites. This is somewhat similar to the conditional knockout strategy. These mutants can be particularly useful when complete removal of the gene, results in early embryonic lethality.

Targeted mutagenesis requires that the gene is known before the phenotype can be ascertained but random mutagenesis relies on identifying phenotypes and then looking for the causative gene. However, finding the disruptions that are responsible for causing the phenotype can be a laborious process. Key data acquired from expression pattern and physical mapping studies, however, can facilitate the identification of the genes affected by speculative or fortuitous mutation. Once the particular gene or genes affected in such mouse models are identified they can provide valuable information critical to the understanding of gene function.

Often mutant mouse strains are initially recognised on the basis of easily identifiable phenotypes that readily set them apart from the many animals being handled in a standard colony or screen. Malformations of hair follicles or the inner ear produce such phenotypes. Mutation of either of these structures is not likely to be lethal, often displaying overt phenotypes and so many mouse models harbouring mutations in these structures have been identified, although animals carrying mutations affecting both hair and inner ear formation are much rarer.

The mouse mutants *yellow submarine (Ysb)* and *light coat and circling (Lcc)* were both isolated on the basis of a recessive circling phenotype and a non-segregating semi-dominant coat colour that made these animals appear more yellow than their wild type littermates (Dong et al., 2002). Both mutant strains are born with the expected Mendelian frequencies from heterozygous matings and homozygotes are

viable and often fertile. In these mutants recessive circling behaviour and head tossing indicates an abnormality in the structure or function of the vestibular component of the inner ear, resulting in balance problems. Further investigation confirmed vestibular abnormalities but also shows that these mice have cochlear defects resulting in deafness.

Coat abnormalities or colour differences are often recognised in screens. The lightness observed here appears to be the result of some disruption in the way a normal agouti coat is formed (Lamoreux et al., 2001). Agouti, a coat with a brownish appearance, is the wild-type coat colour of mice. It results from melanocytes in the base of the growing hair switching melanin production between the two pigments eumelanin (brown to black) and phaeomelanin (red to yellow) as the hair elongates. Two genes are involved in melanin switching: *extension* and *agouti. Extension* encodes the melanocortin 1 receptor (Mc1r) that is able to bind to the α-melanocyte stimulating hormone (α-MSH), thus transducing the signal to produce eumelanin. The agouti protein antagonises the interaction of α-MSH with Mc1r resulting in a switch from eumelanin to phaeomelanin production. Transient expression of *agouti* between days four and six of the hair cycle results in black hairs with the characteristic subapical yellow band. Differences, not only in the relative amounts of pigment and their distribution in single hairs, but also in the proportion of alternatively coloured hairs can give rise to a wide variety of apparent coat colours in the mouse. Normally about eighty percent of the coat is comprised of zigzag type underhairs but auchenes, awls and monotrichs (guard) have a greater influence on the overall coat colour because they are longer and lie over the top of the zigzags.

The *Ysb* mutant arose by the random integration of a *Col2a1-ßgal* reporter transgene in a promoter analysis experiment. There is a complex mutation associated with this mutant which involves at least two integration sites, close to each other, towards the centromeric end of chromosome 3 and a deletion of about twenty kilobases in between these integration sites (Dong et al., 2002).

When stained with X-gal, embryos show a ß-gal expression pattern broader than expected for *Col2a1* and that found in other lines of mice carrying the same transgene. This indicates that the reporter has integrated in a site which is under the influence of an endogenous promoter (Keith Leung and Anna Tang unpublished data). At 9.5 dpc, embryos display additional staining in the hindbrain and a few days later a spotty pattern of expression appears that marks the developing hair follicles. Analysis of embryos homozygous for the *Ysb* allele show that not only are the sensory cells of the inner ear severely disrupted but also the innervation of the ear is abnormal (Keith Leung and Anna Tang unpublished data).

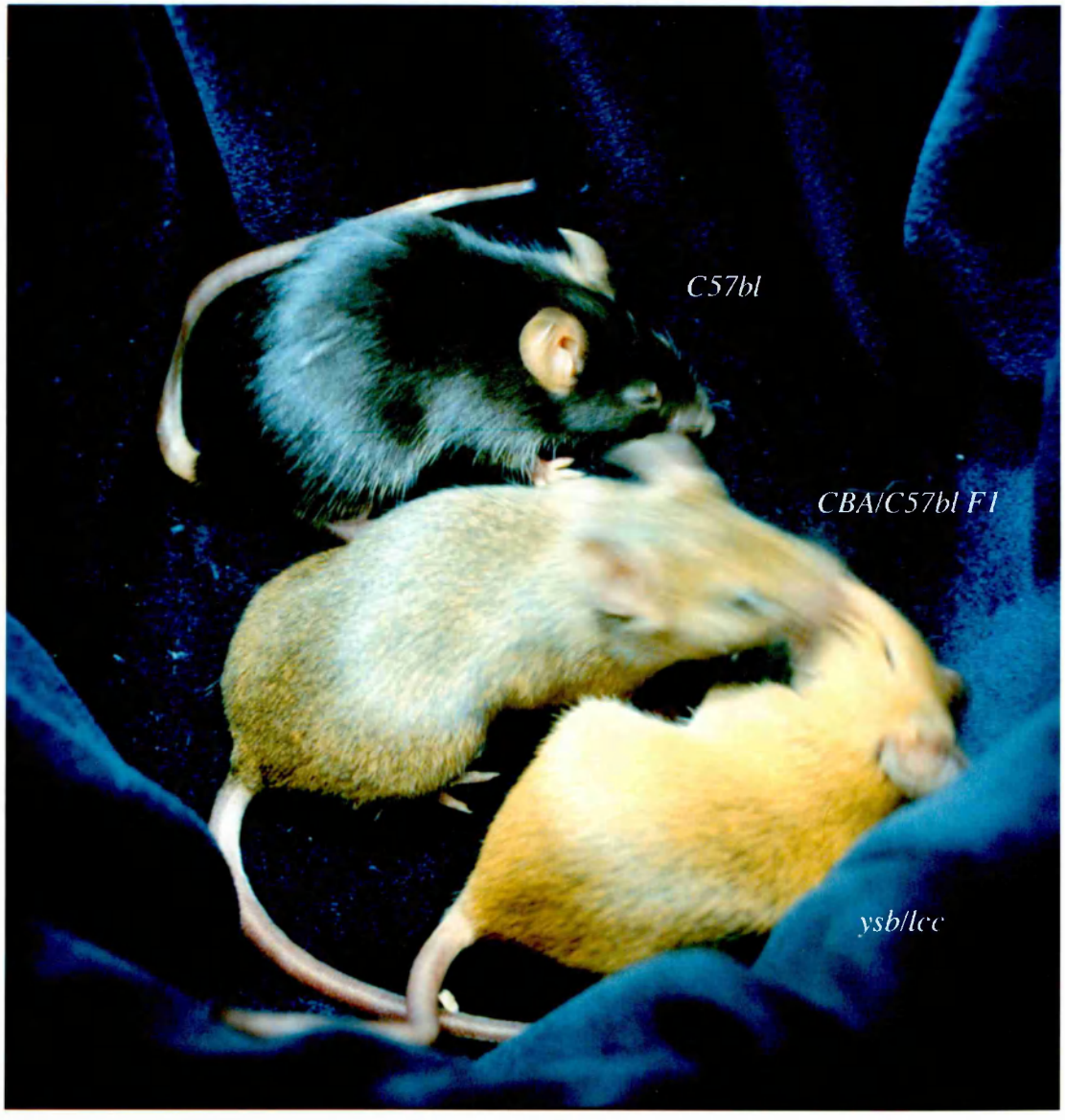
The mouse mutant *Lcc* was isolated at Harwell during an X-ray mutagenesis screen (Lyon et al., 1979) and although the phenotypes exhibited are similar to *Ysb* they appeared to be more severe, i.e. a more yellow coat and more frequent/violent circling. The *Lcc* mutation maps to chromosome three and G-banding analysis reveals that there is an inversion between B and E1 at the centromeric end of this chromosome. Heterozygous *Lcc* mice do not have a circling phenotype, but their coat does appear slightly more yellow when compared to their wild type littermates. (Dong et al., 2002).

Phenotypic similarities between these *Ysb* and *Lcc* mutant indicates that their mutations might be affecting the same gene or genes. When heterozygous animals from each line are bred together, $Lcc^{+/-}/Ysb^{+/-}$ offspring are yellow and circling (Dong et al., 2002) (Fig. 4.1). Since *Lcc* and *Ysb* fail to complement each other in this breeding assay and have also been mapped to the same region of chromosome three it is likely that the mutations are indeed allelic.

Close examination shows that the severity of inner ear defect varies depending on the genotype of the mice. The gross structure of the inner ear can be visualised by filling the lumen of embryonic inner ears, at 16.5 dpc, with latex paint (Amy Kiernan et al., manuscript submitted (after Martin and Swanson, 1993)). Although both *Lcc* and *Ysb* heterozygotes have an essentially normal ear structure, the homozygotes and compound heterozygotes display a range of abnormal inner ear formation (Fig. 4.2a,b,d,e). *Lcc* homozygotes have the most severe phenotype, with many of the vestibular and cochlear components missing (Fig. 4.2e). The form the latex paint acquires in these samples shows that the channels of all three semicircular canals are completely malformed and there are only remnants of the utricular and saccular components where cells that transduce movement into nerve signals should reside. The cochlea, which is responsible for detecting sound, normally forms a coiled structure with one and a half turns. In *Lcc* homozygotes, however, this structure only just manages half a turn and electron microscopy reveals that the normal cellular arrangement of sensory cells is completely missing (Amy Kiernan et al., manuscript submitted). $Lcc^{+/-}/Ysb^{+/-}$ have the next most severe abnormality followed by *Ysb/Ysb* embryos (Fig. 4.2d). The variety in the inner ear structure

Figure 4.1 Light coat colour of circling *Lcc*$^{+/-}$/*Ysb*$^{+/-}$ mice.

Figure showing the typical yellow coat colour of *Lcc*$^{+/-}$/*Ysb*$^{+/-}$ *animals. Matings between the Lcc and Ysb mouse lines produce offspring with a mixed background but compound heterozygotes are readily detected amongst littermates that have normal coat pigmentation. Animals from stock strains show the dramatic colour contrast between black (C57Bl6), agouti (CBA/C57Bl10* F1) *and yellow coats.*


C57bl
CBA/C57bl F1
ysb/lcc

indicates that, although the alleles *Ysb* and *Lcc* probably affect the same gene (or genes), the influence that these disruptions have is not identical.

Chromosomal abnormalities of the *Lcc* and *Ysb* alleles were not the same but there was a region of overlap that is likely to contain the core components critical for proper ear and hair formation. Core components may be influenced by chromosomal abnormalities unique to the respective alleles and hence a range of abnormalities were seen that depended upon the exact genotype.

4.3 Results

In order to explain the phenotypes of the *Ysb* and *Lcc* mutants it is useful to know the actual gene or genes being affected. Since the precise nature of the mutation in these animals is unknown a candidate approach was employed, based upon the approximate physical position of the two mutations and the structures that are affected in these strains. Of all the genes that have been mapped to the centromeric end of chromosome 3 *Sox2* appears to be one of the best candidates in the vicinity of the region disrupted in the *Ysb* and *Lcc* mutants. Apart from physically mapping to the same approximate chromosomal location, expression data shows that *Sox2* is expressed in the same tissues that are affected by the *Lcc* and *Ysb* mutations (Chapter 2).

4.3.1 Breeding with $Sox2^{\beta geo-/+}$ shows that *Sox2* is allelic to both *Ysb* and *Lcc.*

An initial investigation to establish if there is any disruption in the coding region of *Sox2* was carried out by examining genomic DNA from the *Lcc* and *Ysb* mutants by Southern blot. This analysis reveals that the coding region of *Sox2* and the DNA immediately surrounding it did not have any major deletions, insertions or rearrangements (Anna Pelling, unpublished data). Although the coding region of *Sox2* appears intact it still remains a good candidate since the *Lcc* and *Ysb* alleles could be harbouring mutations affecting the regulation of *Sox2* expression rather than the function of the protein. The $Sox2^{\beta geo-/+}$ mutation is precisely known and therefore it can be used to decipher the involvement of the *Lcc* and *Ysb* mutations in the expression of SOX2. In order to test whether the *Lcc* and *Ysb* mutations are allelic to

Sox2, heterozygotes from each strain were bred against $Sox2^{\beta geo-/+}$ and offspring examined for phenotypes including inner ear and coat abnormalities.

Matings were set up between $Sox2^{\beta geo-/+}$ and $Ysb^{-/+}$ or *Lcc/Lcc* animals. Initially 16.5 dpc embryos were harvested in order to examine the formation of their inner ear. These have a normal structure in all embryos except those that carry one of each of the mutant alleles from their parents. $Lcc^{+/-}/Sox2^{\beta geo-/+}$ samples have an abnormality similar in form and comparable in severity to the *Lcc* homozygotes and the $Sox2^{\beta geo-/+}/Ysb^{+/-}$ inner ears have a structure that is most similar to the $Lcc^{+/-}/Ysb^{+/-}$ samples that were examined previously (Fig. 4.2c-f Amy Kiernan unpublished data). Paint fill experiments show that the architecture of the inner ear is severely disrupted apparently due to malformation of vestibular and cochlear sensory epithelium where *Sox2* expression is found. This confirms that *Sox2* is indeed one of the genes, if not the only gene, being affected by the *Lcc* and *Ysb* mutant alleles because the extent and severity of abnormality is not normally observed in animals heterozygous for any of these alleles.

Compound heterozygotes have coat colour abnormalities.

The second characteristic phenotype of *Lcc*, *Ysb* homozygotes or $Lcc^{-/+}/Ysb^{-/+}$ compound heterozygotes is a distinctly yellow coat colour when compared to their littermates. This phenotype might also be attributable to lack of SOX2. To test this, $Sox2^{\beta geo-/+}$ mice were bred against either *Ysb* or *Lcc* heterozygotes and pregnant females were allowed to litter down. The litters produced were weaned and genotyped by PCR from a tail biopsy (materials and methods). No $Lcc^{+/-}/Sox2^{\beta geo-/+}$ pups were ever recovered and it can be assumed that these animals die at or shortly

Figure 1.2 The inner ear structure of mutants carrying *Lcc*, *Ysb* and $Sox2^{\beta geo}$ alleles

The inner ear structure of mutants carrying *Lcc, Ysb and* $Sox2^{\beta geo}$ *alleles was visualized by filling the lumen of the inner ear from 16.5dpc embryos with paint (Amy Kiernan unpublished data). Bar in* ***g*** *is 100μm and applies to* ***a-f*** *also.*

(a). *Embryos carrying a single allele show an inner ear with normal structure, here represented by Ysb/+.*

(b). *Embryos that are homozygous for Ysb display show malformation of the semicircular canals and maculae. The cochlea is also slightly malformed with less turns than the heterozygote.*

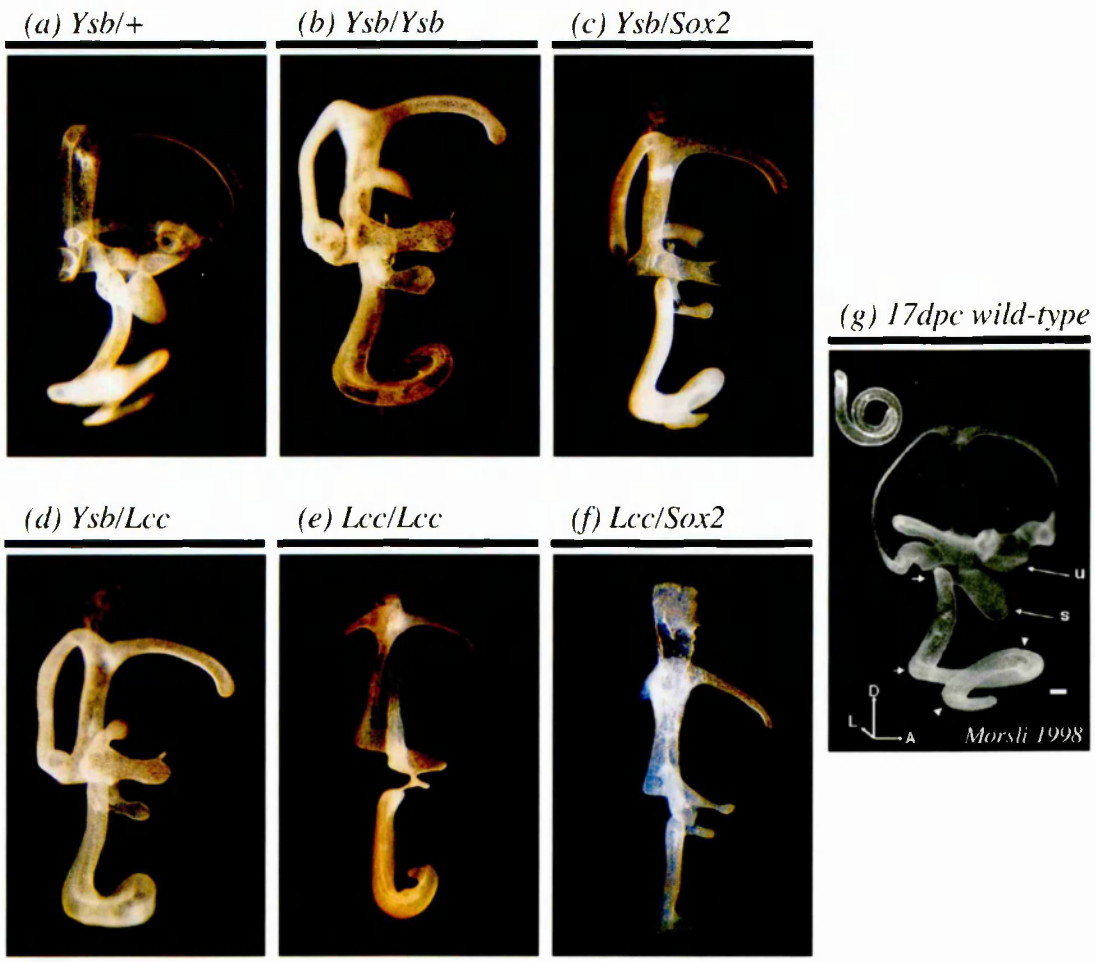
(c). *The compound heterozygote* $Sox2^{\beta geo-/+}/Ysb^{+/-}$ *has a similar but more severe phenotype to that of* ***b.***

(d). *The compound heterozygote* $Ysb^{+/-}/Lcc^{+/-}$ *has an inner ear structure almost identical to* $Sox2^{\beta geo-/+}/Ysb^{+/-}$

(e). *Lcc homozygotes show one of the most severe phenotypes with severe ablation of the vestibular components of the inner ear. The cochlea hardly manages to complete a single turn.*

(f). *A very similar phenotype to* ***e*** *is observed in the inner ears of* $Sox2^{\beta geo-/+}/Lcc^{+/-}$ *embryos.*

(g). *The normal inner ear structure of a wild type 17dpc embryo. The inset is a view of a normal cochlea that makes one and a half turns by this stage of development (Morsli et al., 1998).*


(a) Ysb/+
(b) Ysb/Ysb
(c) Ysb/Sox2
(d) Ysb/Lcc
(e) Lcc/Lcc
(f) Lcc/Sox2
(g) 17dpc wild-type
u
s
D
L
A
Morsli 1998

after birth and are rapidly cannibalised. $Ysb^{+/-}/Sox2^{\beta geo-/+}$ animals, however, do survive and show the circling behaviour expected due to the inner ear abnormalities observed in embryos with the same genotype. $Ysb^{+/-}/Sox2^{\beta geo-/+}$ animals are born at the expected Mendelian frequency indicating that this genotype does not affect viability. Of the compound heterozygotes born, most, but not all display circling behaviour. The coat of these mice is also more yellow than their littermates although a few were pure black. Parents of litters with black pups carry recessive alleles that prevent the production of pheomelanin and so a proportion of black circling pups was expected and represents the animals that are homozygous for this allele. Plucking of hairs and their microscopic examination reveals that, not only is there a difference in ratios of hair type in yellow compound heterozygotes when compared to wild-type littermates, but also in the extent of the pheomelanin band (Keith Leung unpublished data), both of which aspects contribute to the yellow appearance of the coat. These hair phenotypes are similar to those observed in the pure *Lcc* or *Ysb* lines (Dong et al., 2002).

These breeding data suggest that both *Ysb* and *Lcc* are allelic to *Sox2* and since the phenotypes are complemented by a *Sox2* null allele they must cause a reduction or abolition of expression in particular cell types. Such a reduction in the levels of expression is most likely to be a consequence of some disruption in its normal regulation. Promoter sequences responsible for regulating gene expression are generally modular with several distinct DNA sequences required to ensure that the overall pattern of a particular gene's expression is achieved. There can be a considerable physical distance between all the regulatory modules and the gene they control so mutations may only affect gene expression in a subset of tissues at

particular developmental stages. Mutations of this sort can effect gene expression spatially, temporally or in the levels of the gene that is expressed. In this case, both *Lcc* and *Ysb* alleles cause a reduction in the amount of expression but only in a subset of the sites in which it is normally expressed. The effects of the *Lcc* allele appear to be equivalent to that of $Sox2^{\beta geo-/+}$ in compound heterozygotes and so may mediate a complete abolition of *Sox2* in the developing inner ear and hair follicle. The *Ysb* allele produces slightly less severe phenotypes in compound heterozygotes and so either it may not completely abolish *Sox2* expression from this allele or it may cause an abolition over a shorter developmental period that results in a milder phenotype. Mutations that abolish or reduce *Sox2* expression in only a subset of tissues, mediated by *Lcc* and *Ysb* alleles, would explain why homozygous *Lcc* and *Ysb* animals are viable while the homozygous null of *Sox2* is lethal at the peri-implantation stage. *In situ* hybridisation subsequently showed that there was greatly reduced expression of *Sox2* in the otic vesicles of 9.5dpc *Ysb* homozygotes and no *Sox2* RNA or protein could be detected in the inner ear structures of *Lcc* homozygotes at 9.5dpc and 16.5dpc respectively (Amy Kiernan et al., manuscript submitted).

4.3.2 A hair phenotype can be seen in heterozygous $Sox2^{\beta geo-/+}$/C3H mice.

Production of $Lcc/Sox2^{\beta geo-/+}$ or $Ysb/Sox2^{\beta geo-/+}$ mutant embryos and mice shows that *Sox2* is most likely to be allelic to these mutations and that the overall nature of these disruptions is loss of function. Since *Lcc* and *Ysb* both involve complex disruptions of chromosome 3 that could affect more than one gene, however, it is not possible to conclude that these mutations are only affecting *Sox2* or indeed directly affecting *Sox2* on the basis of this evidence. If some aspect of the *Lcc* or *Ysb*

phenotype could be recapitulated in animals in which the *Sox2* mutation is precisely known, then this would suggest that *Sox2* is the most significant, if not the only, gene affected by these alleles. $Sox2^{\beta geo-/-}$ embryos die at the peri-implantation stage but $Sox2^{\beta geo-/+}$ mice are viable and so these animals were examined for semi-dominant phenotypes similar to *Lcc* or *Ysb* heterozygotes that would further implicate *Sox2* as the major gene affected by these mutations and thus having a critical role in both inner ear and hair follicle development.

Of the two predominant phenotypes displayed by animals carrying a single *Lcc* or *Ysb* allele coat colour differences could most readily be detected and measured (Dong et al., 2002). *Lcc* homozygotes clearly have a yellow coat and the heterozygotes, although less yellow, can still be discerned from their wild-type littermates on the basis of coat colour. If no *Sox2* expression in the hair follicle is produced from the *Lcc* allele then it follows that a $Sox2^{\beta geo-/+}$ mouse should also be discernible from its littermates on the basis of coat colour. The fur of $Sox2^{\beta geo-/+}$ mice was therefore examined to see if there is a coat phenotype similar to that observed in the *Lcc* or *Ysb* heterozygotes (Dong et al., 2002).

$Sox2^{\beta geo-/+}$ mice are maintained on two different backgrounds: 129SvEv and MF1. MF1 is albino and so on this background no coat colour phenotype can be seen. 129SvEv mice have an agouti coat colour but carry the *chinchilla* allele (c^{ch}). *Chinchilla* encodes a tyrosinase that interferes with the production of pheomelanin (yellow/red) more than it does with eumelanin (brown/black), resulting in a slightly darker appearance that masks yellow coat phenotypes to some extent (Lamoreux et al., 2001). *Lcc* mutants show a coat colour phenotype when maintained on a C3H/He

background so a congenic breeding regime was therefore set up to change the background of the *Sox2*$^{\beta geo-/+}$/MF1 animals to C3H/He. Changing the genetic background of a mouse is quite a laborious procedure and normally it takes more than ten generations in a congenic breeding scheme to become 99.8% identical across the entire genome, although 94% identity can be achieved in five generations (Silver, 1995). After just two out-breeding generations about 75% of the genome should be isogenic with the new background and at this point *Sox2*$^{\beta geo-/+}$C3H/He/n1f2 mice are discernible from their littermates on the basis of coat colour and are almost equivalent in colour to the *Lcc* heterozygous animals (Fig. 4.3a). This is the first indication that *Sox2* alone is affected in the *Lcc* and *Ysb* mutants (at least for coat colour) since in these *Sox2*$^{\beta geo-/+}$C3H/He/n1f2 animals only 3.5kb including the coding sequence is absent.

4.3.3 Hair samples reveal a *Sox2*$^{\beta geo-/+}$ coat phenotype.

To investigate the hair phenotype more carefully hair plucks were taken from *Sox2*$^{\beta geo-/+}$C3H/He/n1f2 mice and wild-type littermate controls in a manner standard for this type of experiment (materials and methods and (Pennisi et al., 2000b). Briefly three mice for each genotype were selected for analysis and hairs from three mid-dorsal plucks per animal were taken and microscopically examined. In the first comparison, hairs were sorted into types and then relative frequency calculated (Fig. 4.3b). Statistical analysis shows that there is a significant difference in the relative frequencies of hair types present in the samples although there is no difference in the total amount of hairs recovered from the standard plucks. Significant differences are noted only in awl and zigzag hair types between the wild-type and *Sox2*$^{\beta geo-/+}$ samples.

Figure 4.3 The coat phenotype of *Sox2*$^{\beta geo-/+}$C3H/He animals

(a). *The appearance of the coat of a Sox2*$^{\beta geo-/+}$ *mouse on the C3H/He background is distinctly yellow when compared to a wild type littermate.*

(b). *Hair type comparison of Sox2*$^{\beta geo-/+}$ *mice against wild type littermates. Hairs from three mid dorsum plucks from three mice of each genotype were sorted and counted. Statistical analysis of these data, using the unpaired t test, show significant differences in four of the categories analysed.*

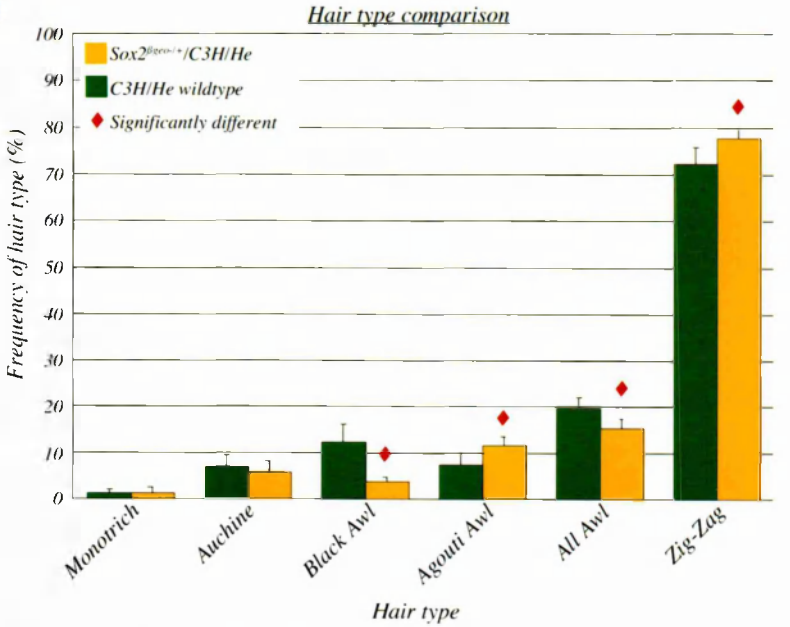
(c). *A Ysb*$^{+/-}$*/Sox2*$^{\beta geo-/+}$ *mouse compared to C57Bl and CBA/C57Bl F1. Coat appears very much more yellow than wild-type counterparts and far more yellow than Sox2*$^{\beta geo-/+}$ *animals.*

(d). *The length of the agouti band was measured in two of the hair types examined. An unpaired t test was then used to examine the null hypothesis that the means of the two experimental populations were equal. Using this test Sox2*$^{\beta geo-/+}$ *awls showed significantly longer eumelanin band with a that those taken from their wild-type littermates (P value of 0.02) but no difference was seen in auchenes.*

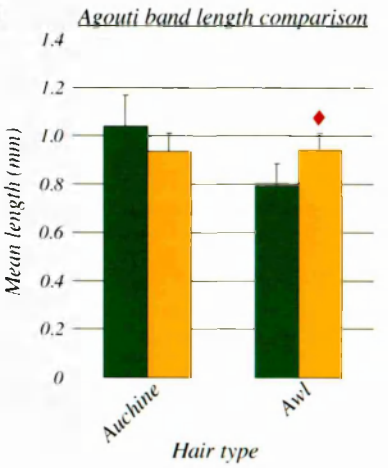
(a)


(c)

(b)


Hair type comparison
Sox2Rex1+/C3H/He
C3H/He wildtype
Significantly different
Frequency of hair type (%)
100
90
80
70
60
50
40
30
20
10
0
Monotrich
Auchine
Black Awl
Agouti Awl
All Awl
Zig-Zag
Hair type

(d)


Agouti band length comparison
Mean length (mm)
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0
Auchine
Awl
Hair type

The relative number of all awls, the most common type of over hair, is significantly lower in samples from *Sox2*$^{\beta geo-/+}$ animals. Sub-dividing the awls into agouti and all-black, it is clear that there is in fact a decrease in the number of black awls and an increase in agouti when compared to wild-type. In a wild-type agouti coat there are usually almost twice as many black awls as agouti but conversely, animals carrying a single *Sox2* null allele have approximately twice as many agouti awls compared to black. The wild-type mouse coat comprises about 70% zigzag underhairs but this number is significantly greater in *Sox2*$^{\beta geo-/+}$ littermates.

Three of the four hair types present in the mouse coat normally display agouti pigmentation: auchenes, awls and zigzags. The second analysis performed examined the length of the agouti band in awls and auchenes. It is notable that, although there is no significant difference in the eumelanin band in auchenes, awls from *Sox2*$^{\beta geo-/+}$ animals have a significantly longer stretch of yellow pigment than littermate controls (Fig. 4.3d). These data indicate that *Sox2*$^{\beta geo-/+}$ animals can be discerned from their wild-type littermates on the basis of coat colour and that the difference in hue of the agouti coat can be attributed to a significant increase in the proportion of agouti hairs formed as well as the extent of the eumelanin band in the most common outer coat hairs. These data did not indicate that there were extra agouti hairs being produced but rather there had been a category switch of some hair types.

These measurable phenotypes of the *Sox2*$^{\beta geo-/+}$ mice correspond to the heterozygous phenotype of both the *Ysb* and *Lcc* coat (Dong et al., 2002). This is a strong indication that *Sox2* is the primary gene contributing to the coat phenotype in these strains and that this due to the loss of function of *Sox2*. Variable phenotypes

are observed with different combinations of the mutant alleles, strongly suggesting that cellular function is dependent upon a particular dose of *Sox2* in these tissues.

4.4 Discussion

The expression pattern of *Sox2* indicates that it has a role in various tissues throughout mouse development but peri-implantation lethality of $Sox2^{\beta geo-/-}$ embryos precludes the use of the homozygous mutation after about 4.5dpc. The production of mice carrying a floxed *Sox2* allele would provide one method by which early lethality may be circumvented allowing the consequences of tissue specific *Sox2* disruption to be observed at later stages. Random or serendipitous mutations may also cause the desired disruption but it is difficult to characterise the specific gene or genes affected in models that are isolated on the basis of an interesting phenotype. By linking these phenotypes with particular genetic defects a great deal may be learnt about the molecular mechanisms underlying normal development.

The mouse mutants *Ysb* and *Lcc* display abnormalities in the development of the sensory components of the inner ear as well as hair specification and pigmentation (Dong et al., 2002). A number of genes, including *Sox2*, share the approximate chromosomal location of these abnormalities but *Sox2* is also expressed in the tissues that are most severely affected by these mutations. In order to test the hypothesis that normal *Sox2* expression is affected by the *Ysb* and *Lcc* alleles a complementation assay was performed. In this assay, mice heterozygous for $Sox2^{\beta geo}$ and either *Lcc* or *Ysb* were created in order to identify any similarities between the effects of these alleles. Compound heterozygotes display both inner ear and hair follicle phenotypes suggesting that they are caused by the loss of function of *Sox2* in these tissues. Close examination of *Ysb* and *Lcc* homozygotes reveals that although *Sox2* expression in the developing inner ear is reduced or absent, chromosomal

abnormalities lie some distance away from the transcribed region of *Sox2*, which is intact.

A defect limited to particular regulatory elements controlling *Sox2* expression would explain why phenotypes are observed only in certain tissues, whereas the complete removal of *Sox2* causes lethality. *Ysb* and *Lcc* alleles may however be directly affecting the expression of several genes, including *Sox2*, with synergy between these defects causing hearing, balance and coat abnormalities. Lack of genetic complementation as well as similar phenotypes in a combination of tissues where *Sox2* is expressed, strongly suggests that this is the key gene affected by these mutations. It remains possible, however, that the disruption of *Sox2* alone may not have any of the same affects reflecting a more minor role in the development of the hair follicle and inner ear. In order to reinforce the evidence that it is *Sox2* alone that is responsible for these phenotypes animals with a far more defined mutation must be examined. $Sox2^{\beta geo-/+}$ mice carry such a mutation and hair specification and pigmentation was chosen as the most accessible phenotype to examine. When compared to wild-type littermates $Sox2^{\beta geo-/+}$ animals have abnormalities comparable to those of *Ysb* and *Lcc* heterozygotes but these are only revealed on a permissive background. This is a strong indication that *Sox2* is important for hair formation, and most likely also for inner ear formation and that the *Ysb* and *Lcc* alleles are a useful tool to examine its role in these and possibly other tissues.

The examination of hairs from the coat of $Sox2^{\beta geo-/+}$ animals reveals that *Sox2* functions in more than one component of murine hair development. Although *Sox2* is expressed during the initial stages of hair follicle formation, induction of follicles

appears to be unaffected in mutant animals. A switch in hair type and pigmentation characteristics is observed in mutants, but both the specification of a particular hair type and the extension of the agouti band manifest themselves in the hair lacking *Sox2*. Of the sites where *Sox2* is expressed in the developing hair follicle, abnormalities in the dermal papilla and dermal sheath cells are most likely to lead to observed phenotypes (Chapter 2).

The dermal papilla is thought to play an important role in follicular development as well as the cycles of hair growth and regression that occur throughout the life of the mouse, however, little is known about the molecular mechanisms underlying its function. This cluster of mesodermally derived cells directs the fate of the adjacent epidermis dictating what type of follicle and consequently hair should be produced. In the hair cycle the dermal papilla does not regress but seems to coordinate the formation of a new follicle, of the appropriate class, from stem cells located in the bulge region (Chapter 2).

Hair pigment is produced by melanocytes at the base of the developing hair shaft and donated, via melanosomes, to adjacent keratinocytes as the hair extends. Dermal papilla cells can excrete the *agouti* signalling peptide that antagonises the interaction of α-MSH with the Mc1r receptor on the surface of melanocytes causing them to switch to the production of the yellow pigment pheomelanin from black eumelanin (see above). When dermal papilla or dermal sheath cells are transplanted *in vivo,* not only are new follicles induced but the hairs produced show growth and pigmentation characteristics typical of the follicles of the donor cells rather than that those of the recipient tissue (Reynolds and Jahoda, 1992; Reynolds et al., 1999). It

follows that any modification of the normal signalling network of the dermal papilla has the capacity to alter signals, such as *agouti,* that influence pigmentation patterns or the form of the hair that is produced.

A single null allele of *Sox2* mediates effects on both the hair type and pigmentation characteristics of the mouse pelage. This indicates that it is likely to play a crucial role, in a dose dependent manner, in the signalling network of dermal papilla cells. If a particular quantity of *Sox2* is required for wild-type function of the dermal papilla then further reduction in the amount of *Sox2* might be expected to have a more pronounced phenotype and indeed the agouti $Sox2^{\beta geo-/+}/Ysb^{-/+}$ animals do have a much more yellow coat colour than with either allele alone (Fig. 4.3d). This mutant, however, may have some residual *Sox2* expression sufficient to obscure a more dramatic function of *Sox2* in the developing hair follicle as it does in the inner ear. *Lcc* homozygotes display the most severe phenotypes in both ear and hair formation and might represent the phenotype of animals with no *Sox2* expression in the hair follicle or the inner ear. If $Sox2^{\beta geo-/+}/Lcc$ animals had been able to survive post-partum they are likely to have displayed an equally (or more) severe coat phenotype.

Exactly how a reduction of SOX2 results in the observed phenotypes remains to be shown but it may have some responsibility in regulating the expression of genes necessary for mediating organising signals from the dermal condensate and later the dermal papilla (Chapter 2). As in other tissues it is possible that the regulatory responsibility for certain genes is shared by different *Sox* proteins. Competition for common regulatory elements could result in differential regulation of a subset of

target genes that in this case might determine the characteristics of the hair being produced. The expression of other *Soxes* in the same cell type might even mask the absolute role of this class of genes in hair follicle development where only one or another is removed. Neither *Sox1* nor *Sox3* expression was ever detected in the developing follicles but interestingly similar but, distinct phenotypes are observed in the pelage of animals carrying alleles with mutations in another *Sox* family member, *Sox18*.

Sox18 is a member of subgroup *F* and is expressed in dermal papilla and dermal sheath cells in a pattern very similar to that of *Sox2* (Pennisi et al., 2000b). Null mutation of *Sox18* shows similar but opposite coat phenotypes to that observed in the $Sox2^{\beta geo-/+}$, *Lcc* or *Ysb* mice. These include a decreased proportion of zigzag hairs and a darkening of overall coat colour due either to the reduction of size of the subapical pheomelanin band or its complete absence (Pennisi et al., 2000a). The authors were surprised at the mild phenotype of these animals, which lack any functional *Sox18,* as they had already established that the classical mouse mutant *ragged* carries missense mutations in the *Sox18* gene which leads to a severe disruption of hair follicle formation and a homozygote that is practically nude (Downes and Koopman, 2001; Pennisi et al., 2000b; Slee, 1957). The conclusion drawn is that *ragged* alleles are contributing to a trans-dominant negative effect mediated by mutant SOX18 proteins rather than haploinsufficiency because of the lack of similar phenotypes observed in the *Sox18* null mutants.

The evidence indicates a synergy between SOX2 and SOX18 as a wild-type coat can only be formed when normal levels of each are expressed. It is tempting to

speculate that the proteins produced from *ragged* alleles are interfering with cellular processes that are normally regulated by a balance between SOX2 and SOX18. Null mutations in either gene have opposite effects on coat formation possibly indicating an involvement in common cellular processes. The absence of either protein results in fairly mild coat phenotypes suggesting that remaining factors have some overlapping function in the maintenance of pelage formation. Acting as a dominant negative *ragged* would perhaps prevent both SOX2 and wild-type SOX18 from performing their normal function, which may be why even heterozygous *ragged* phenotypes are much more severe (Pennisi et al., 2000b; Slee, 1957).

It would be interesting to examine if *ragged* like coat phenotypes could be reproduced in animals carrying combinations of *Sox2* and *Sox18* null alleles. A severe disruption of follicle development in these animals would at least demonstrate a synergy between the two mutations and may provide important information as to whether related *Sox* factors regulate some of the same targets *in vivo*. If found it is possible that similar interactions are also relevant within the regulatory networks of other cell types, such as the gut endoderm, where a similar combination of *Sox* factors is expressed (Kanai-Azuma et al., 2002). There are many tissues where there is overlapping expression of *Sox* genes, the combination of which might determine the identity of particular cells. Although these related factors might regulate completely separate sets of genes a subset of targets may depend upon differential regulation by a group of related factors. The hair follicle could serve as a useful model system in which to investigate *in vivo* interactions between individual *Sox* genes and their targets. It has the additional benefit that slight changes can be readily assayed whilst

more significant disruptions would only cause the loss of this structure without compromising viability.

Together *Sox2*$^{\beta geo}$, *Ysb* and *Lcc* form an allelic series with each of these mutations causing a modification to the normal expression pattern of *Sox2* (and maybe other genes as well). Breeding experiments show that the *Lcc* and *Ysb* alleles cause a loss of function of *Sox2* with the predominant phenotypes observed in the inner ear and the hair follicle. *Ysb/Ysb, Lcc/Lcc* and *Sox2*$^{\beta geo-/+}$*/Ysb*$^{+/-}$ animals are viable, but no *Sox2*$^{\beta geo-/+}$*/Lcc*$^{+/-}$ animals were ever recovered after birth although late gestation embryos were recovered alive and appeared grossly normal. The reason for postnatal lethality of these animals was not determined but is unlikely to be caused by abnormalities of hair follicles or the inner ear. This indicates that there are further sites of *Sox2* expression that are affected by the *Lcc* allele and so it may be used to understand the role of *Sox2* in other tissues of the developing embryo.

Survival of *Lcc* homozygous animals post partum may be explained if there is a certain dose of SOX2 required for viability in whichever tissue is being affected. One explanation for this might be that the *Lcc* allele is affecting a regulatory element responsible for maintaining *Sox2* expression in certain cells after a particular developmental stage. In animals homozygous for the *Lcc* allele, *Sox2* would be expressed at normal levels until the point at which the regulatory mutation means it can no longer be maintained in a particular group of cells. Viability of the whole animal is not lost as enough SOX2 is provided by initial expression. In *Sox2*$^{\beta geo-/+}$*/Lcc*$^{+/-}$ animals only half the dose of *Sox2* is expressed from the outset. This is normally sufficient for viability, however, once *Sox2* expression also ceases from the

Lcc allele, then the amount of SOX2 available in a particular group of cells may drop below a critical threshold and result in the death of these animals.

Expression data has proved key to the identification of *Sox2* as a critical gene disrupted by both the *Lcc* and *Ysb* mutations. Although both of these alleles have large and only partially characterised abnormalities, the link to *Sox2* misregulation makes them a useful tool in the investigation of both inner ear and hair follicle formation as well as to elucidate functional characteristics of *Sox2*. Further investigation may yield additional tissues where *Sox2* expression is disrupted and so facilitate the examination of further aspects of its function. Apart from this, further characterisation of the *Lcc* and *Ysb* alleles may provide clues to the location of regulatory elements that would normally control the expression of *Sox2*. This study highlights the importance of this type of randomly created mutant despite the existence of more deliberate techniques of gene modification such as conditionally null *Sox2* allele. Conditional inactivation of *Sox2* could be use to highlight individual roles in the different components of the hair follicle that exhibit *Sox2* expression. *Cre* constructs utilising versican regulatory elements that direct expression to the dermal papilla could be exploited to extend the current understanding of the role of *Sox2* in these cells.

Chapter 5 Sox gene redundancy addressed by breeding *Sox1*$^{m1+/-}$ and *Sox2*$^{\beta geo-/+}$ together.

5.2 Introduction.

The idea of functional redundancy has arisen largely from the observation of overlapping expression patterns of related genes and their analysis by targeted null mutation. The lack of obvious phenotypes or malformation limited to a subset of expression sites when a single gene is disrupted is often explained as a result of compensation by a related gene (Cheung et al., 2000; Kume et al., 2000; Overton et al., 2002; Pennisi et al., 2000a; Solomon and Fritz, 2002). Such compensation may be abrogated if animals carrying multiple null alleles are created. In these compound null models, phenotypes are often more severe than the additive effects of the individual mutations and this is taken as an indication of functional redundancy between the respective genes. The implication that *Sox1*, *Sox2* and *Sox3* share some functional responsibility may be similarly addressed, *in vivo*. By breeding together mice that carry null mutations in these genes it should be possible to create animals lacking different doses of these genes.

At the time when this experiment was conceived, targeted disruption of all the *SoxB1* genes had been made but only *Sox1* and *Sox2* null alleles were available in viable mouse lines. Although mice carrying null alleles in all three genes may be expected to have the severest phenotype it is important to examine the effect of all the possible combinations of null alleles. An experiment focussing on the effects of combining null alleles of *Sox1* (*Sox1*^{m1}) and *Sox2* (*Sox2*$^{\beta geo}$) was therefore carried out.

Viability was used as the initial crude measure of the effect of these mutations. The two individual mutants were re-examined using this assay before analysing animals carrying combined mutations. If the functions of these genes are unrelated then the addition of their phenotypes should not decrease viability beyond the simple summation of that of *Sox1*$^{m1-/-}$ and *Sox2*$^{\beta geo-/+}$. If there is an increase in mortality above this there are two potential explanations. Firstly, there is functional equivalence between these two genes where it is the overall level of *SoxB1* protein in a particular cell that is important. Reduction of this total level by disrupting increasing numbers of alleles would be expected to result in increasingly pronounced phenotypes. In this model, this effect should be the same regardless of which combination of alleles is removed; the importance is in the absolute dose. The second explanation is that the two proteins are active in the same cells at the same time but their function is qualitatively different. In this case, one might still expect to see a more severe phenotype in these compound mutants but the particular combination of alleles removed is critical. These hypotheses can only be distinguished once the consequence of all possible combinations of mutations has been revealed.

An increase in mortality in the compound mutants would require further investigation in sites of overlapping expression between *Sox1* and *Sox2*. It will be important to analyse phenotypes in these regions to uncover the roles of the *SoxB1* genes in development and would rule out any possibilities of a purely synergistic effect on viability.

5.3 Results

5.3.1 Do single *Sox1*^{m1} or *Sox2*$^{\beta geo}$ alleles affect viability?

Before examining the effect that compound null mutations of *Sox1* or *Sox2* have on mouse viability it was important to analyse what affect these alleles had on their own. It has already been established that the most severe effects of these alleles manifest themselves when there is a homozygous null mutation (Avilion et al., 2003; Malas et al., 2003; Nishiguchi et al., 1998) but it was important to establish the consequence, if any, to animals that carried a single null allele of either of these genes.

The Sox1^{m1} *mutation.*

The *Sox1*^{m1} null mutation has already been described (Malas et al., 2003; Nishiguchi et al., 1998). Litters from heterozygous matings contain the expected Mendelian frequency of null alleles at birth so SOX1 does not appear to be essential for embryonic life. The predominant phenotypes of *Sox1*$^{m1-/-}$ mice are micropthalmia, spontaneous seizures and a high mortality rate in both inbred and hybrid backgrounds (Chapter 1). Breeding data confirmed that, although at weaning age there are less *Sox1*$^{m1-/-}$ mice present than would be expected, a significant number are able to survive well past this age. The numbers of heterozygotes at weaning did not give any indication that a single null allele has any effect on viability up to this age (Table 5.1).

Table 5.1 Tail biopsy data for *Sox1*^{ml} heterozygote matings.

Tail biopsy data was collected over one year from maintenance breeding of Sox1^{ml} *line where both parents are Sox1*$^{ml-/+}$.

Background	Genotype	Weaned (n)	% of Total
129SvEv	+/+	23	27%
(12 litters)	-/+	50	53%
	-/-	13	15%
Total		86	
C57Bl/6	+/+	20	34%
(17 litters)	-/+	31	53%
	-/-	7	12%
Total		58	

The Sox2$^{\beta geo}$ mutation.

As with the *Sox1* null mutation, *Sox2*$^{\beta geo}$ homozygotes showed a far more severe phenotype (Avilion et al., 2003). All *Sox2*$^{\beta geo-/-}$ animals die at the peri-implantation stage and so SOX2 is an essential component for embryonic life. An analysis of the breeding data of the *Sox2*$^{\beta geo}$ mice revealed that there was also a significant detrimental effect of a single null allele. This effect was observed on both the 129/SvEv and MF1 backgrounds although it was greater on the inbred background. Tail biopsy data, from over one hundred litters, created by mating *Sox2*$^{\beta geo-/+}$ males against wild-type females, indicated that at weaning age (3-4 weeks post-partum) there were significantly less *Sox2*$^{\beta geo-/+}$ animals present than expected, assuming normal Mendelian inheritance. When numbers and genotypes were examined, forty or twenty-nine percent less *Sox2*$^{\beta geo-/+}$ mice were present than would have been expected on the 129SvEv or MF1 backgrounds respectively (Table 5.2).

In order to establish when *Sox2*$^{\beta geo-/+}$ animals were dying, carcasses of pre-weaned mice found dead in cages were collected from maintenance breeding of the *Sox2*$^{\beta geo-/+}$ lines and genotyped. Analysis revealed that almost 74% (n=23) of these animals were *Sox2*$^{\beta geo-/+}$ which could entirely account for the discrepancy seen at three to four weeks post partum. Animals tended to die either soon after birth or after about three weeks. No gross abnormalities were observed in the younger group of animals but three week old *Sox2*$^{\beta geo-/+}$ pups were often, although not always, much smaller than their wild-type littermates.

Table 5.2 Tail biopsy data for *Sox2*$^{\beta geo}$ maintainence breeding.

Tail biopsy data was collected over one year from maintenance breeding of Sox2$^{\beta geo}$ *line. Father is always Sox2*$^{\beta geo-/+}$ *and mother wild-type.*

The significance of χ^2 *analysis with five degrees of freedom is P< 0.05 if* χ^2 *is greater than 11.07 and P<0.01 if the value is greater than 15.09. *denotes estimates made by normalising expected figures against the observed wild-type samples.*

% Missing is calculated on the basis that wild-type numbers represent the norm.

Background	Genotype	Weaned (n)	% of Total	% Missing*
129SvEv	+/+	166	62.4%	
(58 litters)	-/+	100	37.6%	40%
Total		266		
χ^2		17.5/43*		
MF1	+/+	259	58.3%	
(69 litters)	-/+	185	41.7%	29%
Total		444		
χ^2		12.7/29*		

5.3.2 Production of *Sox1*^{ml}/*Sox2*$^{\beta geo}$ compound null mutants.

Breeding data confirms the effects on viability to animals carrying null alleles of either *Sox1* or *Sox2*. A mating regime was devised in order to analyse viability in animals carrying combinations of these two null alleles. Mice heterozygous for the null mutation of either *Sox1* or *Sox2* were mated together to produce animals that were heterozygous for each of these null alleles. A second round of breeding was then required to create litters that might carry all the possible genotypes. Offspring from this second round of matings were examined at a variety of stages to establish any effect that these null alleles might have on viability. Matings were selected to control against any effect of genetic background and fell into the two main categories; inbred (129/SvEv) or hybrid (129/SvEv X MF1) background. The observed consequences of the compound null mutations were the same on both backgrounds, although possibly slightly more severe on the inbred background. The majority of data was collected from mice on the hybrid background that generally produced more reliable litter sizes and therefore numbers for statistical analysis.

Statistical analysis of the breeding data.

Apart from background differences, breeding data falls into three main categories. First, all the embryos harvested at different stages in development. Second, all the (recovered) animals that were born, including those that died before weaning and, lastly, those mice that survived for at least four weeks and were then weaned from their mother.

Expected overall Mendelian ratios were calculated allowing for the fact that *Sox2*$^{\beta geo-/-}$ is always lethal at peri-implantation stages and that litters obtained were

derived from a variety of parental genotypes. There is also no linkage between any of the alleles being examined.

Embryonic data.

The effects of the two null alleles were examined during embryogenesis. Embryos were harvested at all stages after mid-gestation until birth and were then genotyped. These data indicate that all expected genotypes are represented throughout embryonic development and that no excessive lethality was noted (Table 5.3). Chi squared analysis (χ^2) confirmed that there is no significant variation from the expected frequencies of genotypes during embryogenesis whether the data is examined as a whole or as subdivisions based on age or background (data not shown). It was also noted that embryos harvested at all stages showed signs of life (heartbeat and reflex movement at later stages).

Weaned animals breeding data.

Tail biopsy data from all breedings carried out in this experiment were analysed to gain an insight into the genotypes that were able to survive up until weaning age (three to four weeks post partum). Statistical analysis of the total weaned animals, using Chi squared, indicates that overall there is a very significant deviation from the expected Mendelian ratios (Table 5.3). Analysis reveals that the deviation from expected frequencies is partly due to the complete absence of $Sox1^{ml-/-}/Sox2^{\beta geo-/+}$, but it is not obvious if the numbers of animals with other genotypes are also significantly different.

Table 5.3 Observed numbers and statistical analysis of MF1 offspring the *Sox1*^{m1}, *Sox2*$^{\beta geo}$ breeding experiment.

Breeding data is split into two categories: 'Embryos' recovered up to 17.5dpc and 'Weaned' animals that had a tail tip biopsy.

The significance of χ^2 analysis with five degrees of freedom is P< 0.05 if χ^2 is greater than 11.07 and P<0.01 if the value is greater than 15.09.

Z_0 is a measure (according to binomial distribution) of the null hypothesis that observed equals expected. P<0.05 if Z_0 lies outside the interval ±1.96 and P<0.01 if Z_0 lies outside the interval ±2.58.

**denotes estimates made by normalising expected figures against the observed wild-type samples.*

	Embryos			Weaned				
Genotype	**Obs. (n)**	**Exp. (n)**	**Zo**	**Obs. (n)**	**Exp. (n)**	**Zo**	***Exp. (n)**	***Zo**
Wild-type	**36**	37	-0.2	**81**	62	2.58	81	0
$Sox1^{m1-/+}/Sox2^{wt/wt}$	**72**	75	-0.4	**133**	105	3.16	137	-0.4
$Sox1^{wt/wt}/Sox2^{\beta geo-/+}$	**73**	57	2.33	**72**	71	0.17	92	-2.3
$Sox1^{m1-/+}/Sox2^{\beta geo-/+}$	**117**	114	0.37	**132**	122	1.08	159	-2.5
$Sox1^{m1-/-}/ Sox2^{wt/wt}$	**32**	37	-0.9	**35**	43	-1.2	56	-2.9
$Sox1^{m1-/-}/Sox2^{\beta geo-/+}$	**47**	57	-1.4	**0**	51	-7.6	67	-8.7
Total (n)	**377**			**453**				
χ^2	**7.3**			**66.8/*83.5**				

Statistical analysis - explanation

In this experiment, statistical analysis relies upon producing an expected value of animals with a particular genotype. These values are normally calculated by multiplying the proportion of animals expected, according to Mendelian genetics, by the total number of samples recovered and is useful when examining a shift from one genotype to another. In this experiment the total number of samples at weaning is an underestimate as there is a proportion of animals that die, due to their genotype with respect to $Sox1^{m1}$ or $Sox2^{\beta geo}$. Many of these animals are lost as they are cannibalised before they can be counted and genotyped. In order to estimate the actual number of animals that should have survived to weaning, if there was no detrimental effect of a particular genotype, an expected total was extrapolated by multiplying up from the number of wild-type pups. A certain proportion of wild-type animals will die but this number ought to be equal across the respective genotypes and therefore this method gives a good baseline estimate.

By making these estimates for all weaning data and comparing them to actual values it can be shown that significantly fewer animals of all genotypes, except $Sox1^{m1-/+}$ and wild-type, survived to weaning (Table 5.3). However, animals with the $Sox1^{m1-/-}/Sox2^{\beta geo-/+}$ genotype are the only ones that never survive past weaning age.

Death before weaning age.

Breeding data shows that there is a major discrepancy between animals expected to be present at weaning and the actual numbers observed, although all expected genotypes were present during embryogenesis. An analysis of animals that had died within four weeks post partum was therefore performed. As many dead

Figure 5.1 The longest surviving *Sox1*$^{m1-/-}$/*Sox2*$^{\beta geo-/+}$ animals do not thrive.

A few mice carrying the severest expected genotype were on the hybrid background and survived to about three weeks old. These animas never thrived or survived being weaned despite being indistinguishable from their littermates at birth.


Sox1m1-/+/Sox2wt/wt
Sox1m1-/-/Sox2βgeo-/+

pups as possible were collected and genotyped by PCR. It was often difficult to recover all the carcasses before the mother had eaten them and so it was assumed that these animals only represented a proportion of those that died before weaning.

The majority of all animals found dead were up to one day old (58%, n=58). There was another cluster of deaths at about three weeks (17% 19-23d, n=17). Of the twenty-one *Sox1*$^{m1-/-}$/*Sox2*$^{\beta geo-/+}$ animals found dead only five survived past three days old and animals of this genotype were the only ones that all died before weaning age. Although there was no discernible difference in size at birth surviving *Sox1*$^{m1-/-}$/*Sox2*$^{\beta geo-/+}$ animals were always a fraction of the size and weight of their littermates at three weeks (Fig.5.1). No obvious physical or behavioural attributes were observed that could explain this difference. Other animals that were either *Sox1*$^{m1-/-}$ or *Sox2*$^{\beta geo-/+}$, although significantly more likely to die before weaning than wild-type littermates, did not differ greatly in size and were usually able to survive past four weeks old.

By adding together all born animals (whether dead or alive) an estimate of the total born animals can be made ('*Exp.' Table 5.4). This indicates that only about half of all dead pups were recovered (difference between born 'total' and born '*total') and that most of this difference is due to the absence of *Sox1*$^{m1-/-}$/*Sox2*$^{\beta geo-/+}$ animals. Since almost all of the older carcasses were genotyped it was assumed that most of the 'dissappeared' pups were lost soon after birth and entirely eaten.

Summary of breeding data findings

Examination of all the breeding data reveals that there is a roughly equivalent detrimental effect to carrying a homozygous null *Sox1* mutation or a single null *Sox2* allele. Although embryonic viability is not affected in mice with these genotypes

Table 5.4 Numbers of genotyped carcases and summary of animals born during the *Sox1*^{m1}, *Sox2*$^{\beta geo}$ breeding experiment.

Breeding data is split into two categories: 'Dead PW' are all the post partum carcasses recovered before weaning and 'Born' animals which is the sum of all weaned animals plus post partum carcasses recovered before weaning.

The significance of χ^2 *analysis with five degrees of freedom is* $P < 0.05$ *if* χ^2 *is greater than 11.07 and* $P<0.01$ *if the value is greater than 15.09.*

Z_0 *is a measure (according to binomial distribution) of the null hypothesis that observed equals expected.* $P<0.05$ *if* Z_0 *lies outside the interval* ± 1.96 *and* $P<0.01$ *if* Z_0 *lies outside the interval* ± 2.58.

**denotes estimates made by normalising expected figures against the observed wild-type samples.*

	Dead PW (all)					Born				
Genotype	**Obs. (n)**	**Exp. (n)**	**Zo**	***Exp. (n)**	***Zo**	**Obs. (n)**	**Exp. (n)**	**Zo**	***Exp. (n)**	***Zo**
Wild-type	**7**	12	-1.62	7	0	**88**	74	1.69	88	0
$Sox1^{m1-/+}/Sox2^{wt/wt}$	**16**	22	-1.43	12	1.14	**149**	127	2.27	150	-0.1
$Sox1^{wt/wt}/Sox2^{\beta geo-/+}$	**17**	16	0.37	9	2.96	**89**	86	0.31	102	-1.4
$Sox1^{m1-/+}/Sox2^{\beta geo-/+}$	**23**	29	-1.23	16	1.99	**155**	150	0.44	178	-2
$Sox1^{m1-/-}/ Sox2^{wt/wt}$	**17**	10	2.52	5	5.21	**52**	52	-0	62	-1.3
$Sox1^{m1-/-}/Sox2^{\beta geo-/+}$	**21**	13	2.41	7	5.41	**21**	64	-5.7	76	-6.7
Total (n)	**101**					**554**			655	
χ^2	**15.9/*61.3**					**35.6/*45.7**				

animals are 30-40% more likely to die before weaning age than their wild-type (or *Sox1*$^{m1-/+}$) littermates. Also *Sox1*$^{m1-/-}$/*Sox2*$^{\beta geo+/-}$ compound mutations are not detrimental to embryonic life but are most commonly lethal soon after birth. The few *Sox1*$^{m1-/-}$/*Sox2*$^{\beta geo-/+}$ mice able to survive up until about three weeks post-partum were all on the hybrid background, clearly not thriving and never survived weaning. These data confirm that there is a significant disadvantage in carrying two mutant copies of *Sox1* and one of *Sox2* that exceeds the additive effects of each allele alone.

5.3.3 Search for a phenotype.

As the majority of *Sox1*$^{m1-/-}$/*Sox2*$^{\beta geo+/-}$ animals die at or around birth it is likely that the combined effects of these mutations are exerting themselves during development when both these genes are mainly expressed. The uterus is an extremely supportive environment and embryos with quite major abnormalities are able to survive whilst inside. In the case of the *Shh* null mutation embryos develop with cyclopia and extensive skeletal abnormalities but do not die until birth commonly followed by the rapid cannibalisation of stillborn pups (Chiang et al., 1996). It is for this reason that *Sox1*$^{m1-/-}$/*Sox2*$^{\beta geo-/+}$ embryos, at various stages, were examined in an attempt to identify any abnormality that might be causing premature death post-partum.

Why look at the CNS?

Of the three *SoxB1* genes, *Sox1* has the most limited expression pattern being restricted almost exclusively to the developing CNS, with *in vitro* experiments confirming a role in the early steps of nervous system development (Li et al., 1998; Wichterle et al., 2002). Development of the vertebrate CNS relies upon the highly

orchestrated production of a wide array of neuronal and glial cell types from multi-potential progenitor cells. The first sign of neural progenitors is seen with the formation of the neural plate at around 7.5dpc that subsequently folds to form a tube like structure surrounding a central ventricle. After midgestation, cell proliferation and differentiation result in a dramatic elaboration of the basic neural tube structure in a process that is broadly similar from beginning to end. Throughout development CNS precursor cells proliferate in the ventricular zone that lines the lumen of the developing brain and spinal cord. Once these cells have undergone their final mitosis they move away from the ventricular zone and begin to differentiate into the class of cell appropriate to their temporal and spatial position. Early differentiating neurons are therefore found in the sub-ventricular zone but continue to migrate outward achieving terminal differentiation in more lateral regions.

Whilst many regulatory genes are expressed in the progenitors of particular CNS structures, members of the *Sox* family cross these boundaries and instead seem limited by more general criteria such as the level of differentiation or proliferative capacity (Cheung et al., 2000; Graham et al., 2003). *SoxB1* genes have significantly overlapping expression patterns in the CNS that predominantly mark cells in the ventricular and sub-ventricular zones throughout development (Chapter 1), (Cheung et al., 2000; Collignon, 1992; Wood and Episkopou, 1999). Lethality of $Sox1^{m1-/-}/Sox2^{\beta geo-/+}$ mice may therefore be caused by abnormalities in these populations leading to more widespread effects upon the developing CNS.

The majority of cells in the cortex and striatum of the forebrain develop from two adjacent telencephalic fields. Progenitor cells lining the ventral side of the lateral

ventricle give rise to striatal neurons (including the dorsal striatum, nucleus accumbens and the olfactory tubercle). *Gsh2* is necessary for the maintenance of striatal progenitors and in its absence striatal development is severely disrupted and the lateral ganglionic eminence is dramatically reduced in size (Szucsic 1997, Toresson 2000). The dorsal ventricular zone gives rise to cells that migrate radially eventually forming six cortical layers in a characteristically inside out fashion. Cells born at different times between 12dpc and 16dpc normally contribute to progressively more superficial layers. The *Pax6* mouse mutant, *Small eye* ($Pax6^{Sey}$), produces no functional *Pax6* causing the death of homozygotes soon after birth. *Pax6* is highly expressed in cortical progenitors and its absence (in $Pax6^{sey/sey}$ embryos) causes a partial redistribution of metaphase cells due to an acceleration of cortical development (Estivill-Torrus et al., 2002). Transiently induced chemical lesions of cortical precursors results in thinning of the cortical plate but this is partially compensated for with the normal cortical arrangement and connections still being produced (Gillies 1993).

No explicit abnormalities have been described in the CNS of $Sox2^{\beta geo-/+}$ embryos. The $Sox1^{m1-/-}$ mutation however causes a specific lesion in the olfactory (Piriform) cortex despite a far more widespread expression pattern. This has been shown to be the cause of spontaneous seizures and secondary neural defects in these animals between four and six weeks of age (Chapter 1), (Malas et al., 2003). The absence of more extensive malformation is thought to be due to functional compensation by other *SoxB1* genes. It is possible that this striatal lesion is exacerbated in animals that also carry a null *Sox2* allele or that other malformations arising from the disruption of progenitors are produced. A gross survey of whole

embryos and brains was initially performed but attention was focussed on examining developing forebrain regions after midgestation.

Gross examination of whole embryos

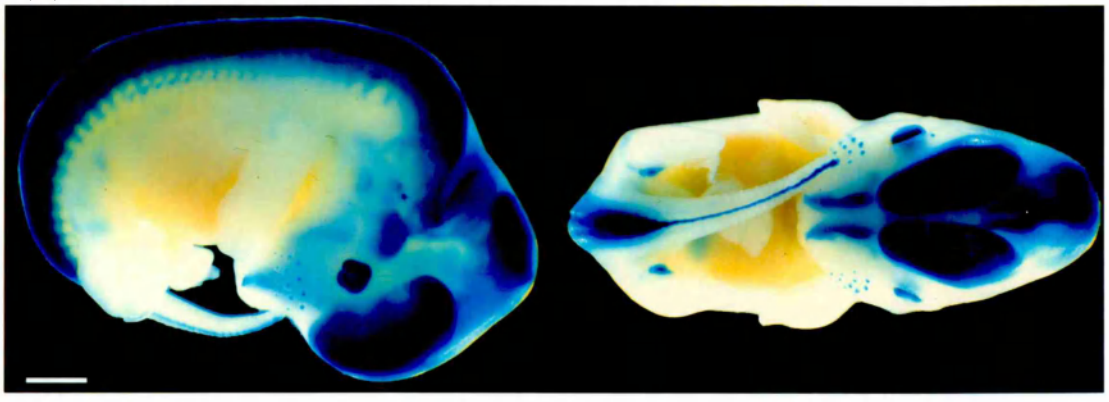
Sox1$^{m1-/-}$/*Sox2*$^{\beta geo+/-}$ embryos were taken from the hybrid line and were examined alongside littermate controls. Where possible, litters for analysis were chosen containing all the expected genotypes but often did not contain a representative of each. For the preliminary analysis it was most important to compare *Sox1*$^{m1-/-}$/*Sox2*$^{\beta geo-/+}$ samples (usually two) with a wild-type or *Sox1*$^{m1-/+}$/*Sox2*$^{wt/wt}$ normal control and a homozygous *Sox1* null as this mutation is known to cause specific CNS lesions and eye defects.

A gross examination of all embryos harvested at different developmental stages was carried out in order to identify any major defects that might result from the combination of disrupted alleles in *Sox1*$^{m1-/-}$/*Sox2*$^{\beta geo-/+}$ animals. Embryos of all genotypes appeared normal at all stages examined and could not be identified on the basis of their appearance. In order to aid the gross analysis some embryos were stained to detect *βgeo* activity conferred by the *Sox2*$^{\beta geo}$ allele. This identifies all of the structures that should have been expressing *Sox2* and allow a comparison between genotypes that carry the *Sox2*$^{\beta geo}$ allele. Again there were no gross differences in the X-gal expression pattern of any of the embryos examined that could be attributed to the *Sox1*$^{m1-/-}$/*Sox2*$^{\beta geo-/+}$ phenotype although micropthalmia, due to the absence of SOX1, was observed (Fig. 5.2 and data not shown). Stained embryos were also wax embedded and sectioned but no obvious differences could be determined and so a more thorough examination of embryonic CNS was undertaken.

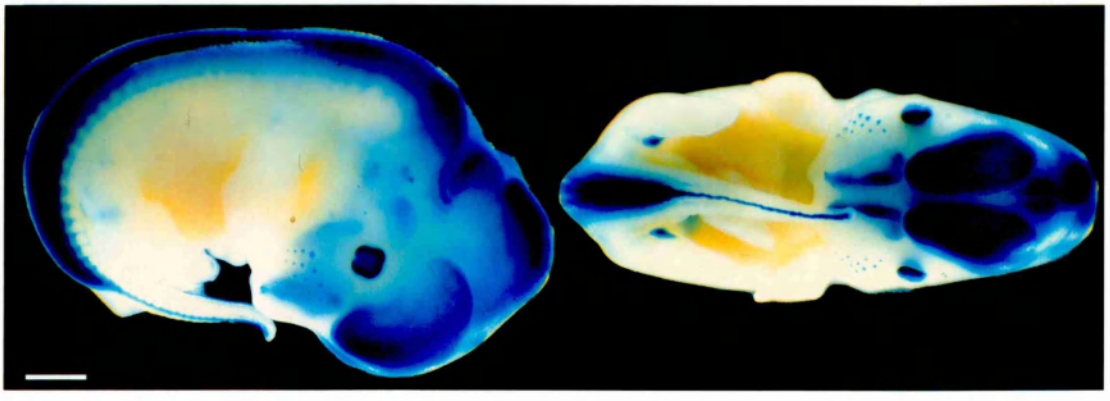
Figure 5.2 12.5dpc X-gal stained embryos from the *Sox1*^{ml}, *Sox2*$^{\beta geo}$ *breeding experiment.*

Littermates on a hybrid background were harvested and subjected to X-gal staining (materials and methods). No gross differences were observed in wholemount except a slightly smaller eye in embryos null for Sox1. Slight differences in staining were more likely to be an artefact of the staining procedure. Bar is 1mm.

(a) *Sox1*$^{m1-/+}$/*Sox2*$^{\beta geo-/+}$



(b) *Sox1*$^{m1-/-}$/*Sox2*$^{\beta geo-/+}$



Gross histological analysis of embryonic CNS.

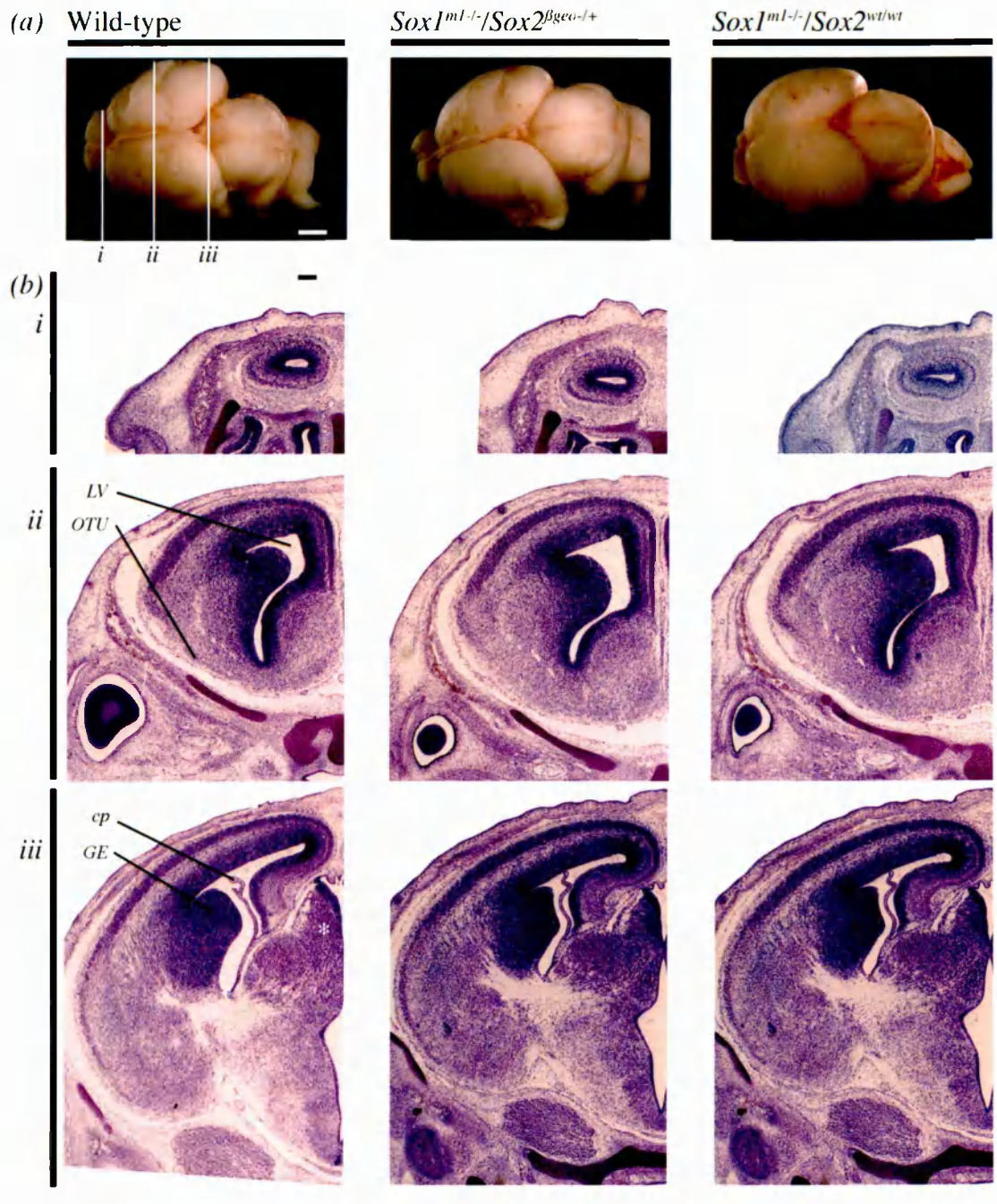
Brains removed from the cranium of 16.5dpc embryos did not differ greatly between genotypes in size or shape and there were no gross abnormalities observed with all the major structures present (Fig. 5.3a). Samples were wax embedded and serial sectioned in order to allow closer histological and immunohistochemical analysis (materials and methods). Series of sections were stained using either haematoxylin and eosin or Cresol violet (to detect Nissl substance in neurons). These were then photographed under high power magnification and comparable sections compared side by side. In this analysis particular attention was paid to the main CNS structures and areas that derive from regions where there is expression of the *SoxB1* genes (Fig. 5.3b).

It is possible that the reduction in *SoxB1* genes in $Sox1^{m1-/-}/Sox2^{\beta geo-/+}$ embryos disrupts telencephalic progenitors and might produce thinning or disruption of the ventricular cell layers or structures derived from these areas. No lesions or atrophy of the ventricular zones can be seen in $Sox1^{m1-/-}/Sox2^{\beta geo-/+}$ samples although this is where there is the greatest expression of *Sox1, Sox2* and *Sox3* in the developing CNS. A migration of ventricular progenitors also gives rise to the nucleus accumbens and the olfactory tubercle that is disrupted in *Sox1* null embryos. This specific lesion is difficult to visualise using standard histological techniques but does not appear to be more severely disrupted in animals carrying an additional *Sox2* null allele. Comparison of all sections (shown) indicated that there are no gross differences in any of the forebrain structures at 16.5dpc. Overall, comparable sections appeared

Figure 5.3 Gross examination of the CNS in embryos of the major genotypes being examined.

(a). *Whole brains were removed from the cranium of 16.5dpc embryos and compared side by side but no gross abnormalities could be observed. Samples were also processed for histological analysis. The position of coronal sections shown are depicted on the wild-type brain sample. Bar is 1mm.*

(b). *First row (plane* ***i****) cuts through the olfactory bulbs. The second series of sections (plane* ***ii****) shows the lateral ventricle (LV) at the level of the olfactory tubercle (OTU) that is disrupted in animals carrying the* $Sox1^{m1-/-}$ *genotype (micropthalmia is also noticeable in these samples). The final set of sections (plane* ***iii****) cuts through the posterior part of the lateral ventricle showing the choroid plexus (cp), the beginnings of the developing hippocampus (asterisk) and the ganglionic eminence (GE) which is made up of the ventricular and sub ventricular zones. Bar is 100μm for all panes in this section.*


(a)
Wild-type
Sox1m1-/-/Sox2βgeo-/+
Sox1m1-/-/Sox2wt/wt
i
ii
iii
(b)
i
ii
LV
OTU
iii
cp
GE

extremely similar and there was no obvious difference in the dimensions or organisation of the CNS discernable by this type of analysis.

Immunohistochemical analysis

Histological examination of serial sections of brains did not reveal any obvious difference between samples carrying different combinations of null alleles. It was then necessary to focus any search by analysing a series of comparable serial sections with molecular markers that would detect more subtle changes in CNS structure. There is an extraordinary number of markers that could indicate anomalies in brain samples, but the initial choice was made as general as possible. Antibodies to detect SOX1, SOX2 and SOX3 were employed in order to indicate precisely regions of the CNS where they are expressed. When used in combination on serial sections a clear picture could be obtained as to which cells have overlapping expression of these genes and perhaps give an indication of the regions of the developing CNS that might be most affected in the $Sox1^{ml-/-}/Sox2^{\beta geo+/-}$ animals.

Sox protein expression in 11.5dpc CNS.

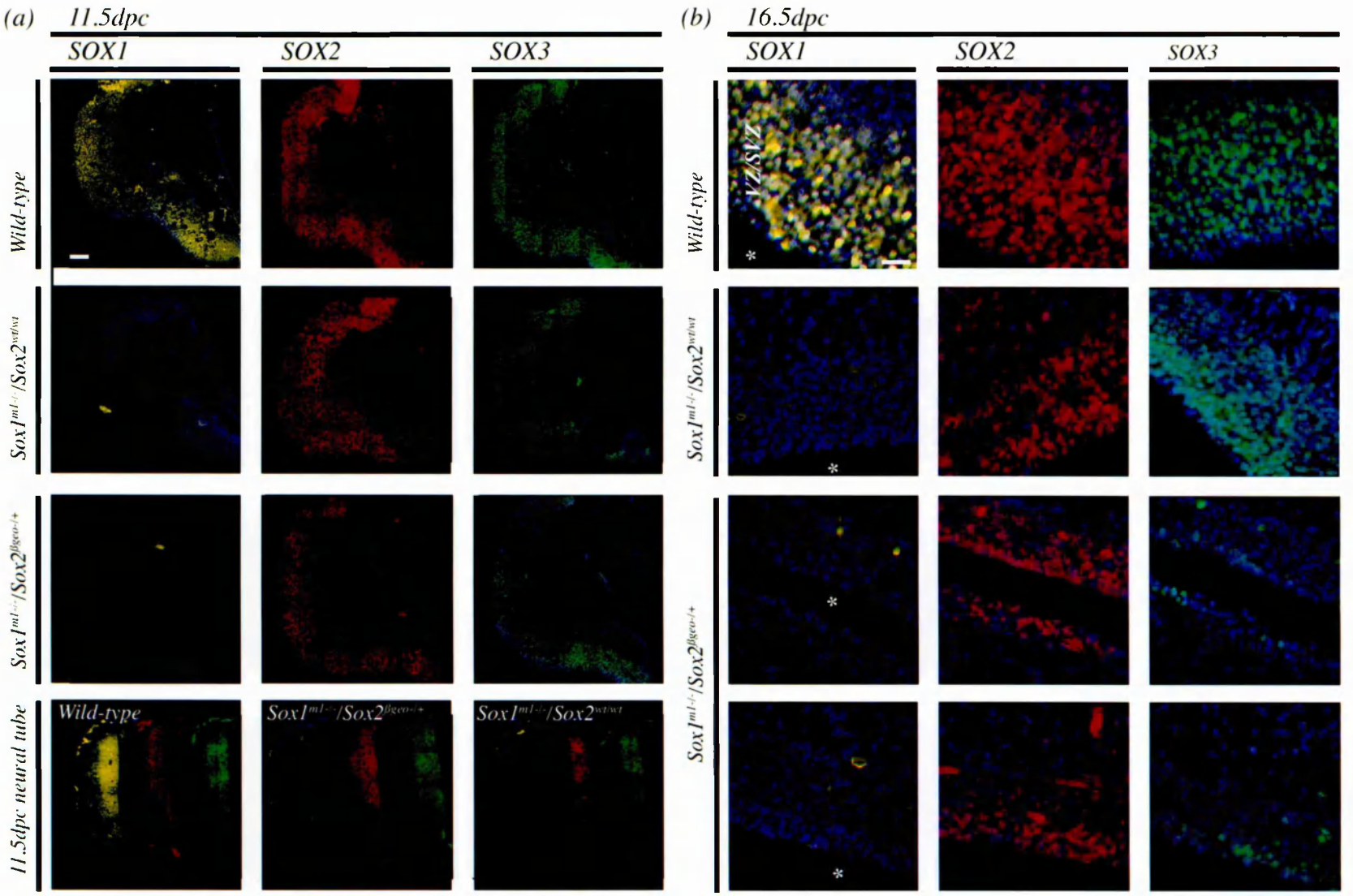
In wild-type 11.5dpc samples the expression of SOX1, SOX2 and SOX3 is very similar and marks cells in the ventricular and sub-ventricular zones, lining the lumen of the developing CNS (Fig. 5.4a). In samples with the $Sox1^{ml-/-}$ genotype, no SOX1 protein was detected and the pattern of the other markers appeared unchanged although generally the remaining SOXB1 proteins were more difficult to detect.

Figure 5.4 SOX B1 protein expression in the ventricular zones and neural tube of 11.5dpc and 16.5dpc embryos.

(a). *11.5dpc samples: Antibodies were used to detect SOX1, SOX2 and SOX3 in coronal brain sections from the three major genotypes being examined. Protein was detected in serial sections where possible and in comparable sections where not. Bar is 100μm.*

(b). *16.5dpc samples: Antibodies were used to detect SOX1, SOX2 and SOX3 in sagittal brain sections from the three major genotypes being examined. Protein was detected in anterior and posterior regions of the third ventricle in serial sections where possible and in comparable sections where not. Asterisk marks the lumen of the lateral ventricle. VZ/SVZ indicates the region of the ventricular and sub ventricular zones where SOX1, SOX2 and SOX3 staining is normally seen in wildtype samples. In samples from* $Sox1^{m1-/-}$ */*$Sox2^{\beta geo-/+}$ *embryos SOX2 and SOX3 staining is detected but the number of positive cells and the intensity of staining consistently appears to be reduced despite these layers appearing similar under conventional histological examination.*

SOX1 appears as pseudo colour yellow. SOX2 signal appears red and SOX3 is pseudo coloured green. Bar is 100μm.


(a)
11.5dpc
SOX1
SOX2
SOX3
Wild-type
Sox1m1-/-/Sox2wt/wt
Sox1m1-/-/Sox2βgeo-/+
11.5dpc neural tube
Wild-type
Sox1m1-/-/Sox2βgeo-/+
Sox1m1-/-/Sox2wt/wt
(b)
16.5dpc
SOX1
SOX2
SOX3
Wild-type
VZ/SVZ
Sox1m1-/-/Sox2wt/wt
Sox1m1-/-/Sox2βgeo-/+

Immunohistochemical analysis of SOXB1 proteins in the ventricular layers of 16.5dpc samples confirmed their continued expression in the ventricular and subventricular zone of the lateral ventricles (Fig. 5.4b). In wild-type samples, protein expression is robust and can be seen in a similar pattern in *Sox1*$^{m1-/-}$ samples except that there is no SOX1 present. Positive cells are normally detected throughout the ventricular zones but generally staining is strongest in cells closest to the lumen. In *Sox1*$^{m1-/-}$/*Sox2*$^{\beta geo-/+}$ samples, although SOX2 and SOX3 are detected, the number of positive cells and the intensity of staining consistently appears to be reduced despite these layers appearing similar under conventional histological examination (Fig. 5.3b). Again antibody detection with the SOX2 and SOX3 antibodies was more difficult in samples carrying the Sox1$^{m1-/-}$ mutation and might indicate a disruption of auto/cross-regulation of these genes resulting in a reduction in dose over and above that attributable to the null alleles alone.

Histological examination of embryos carrying different combinations of null alleles for *Sox1* and *Sox2* did not reveal any gross abnormality in the embryonic CNS, however, there did appear to be an alteration in SOX protein expression that only occurred in animals with the most severe genotype (*Sox1*$^{m1-/-}$/*Sox2*$^{\beta geo-/+}$). The difference in *Sox* protein expression was noticeable at later stages in development and suggests that a reduction in the dose of *Sox1* and *Sox2* genes disrupts proliferating and/or early differentiating cells in the ventricular zones. In these samples, less cells stained positive for SOXB1 proteins which may indicate that cells start to differentiate earlier or that cell cycle dynamics and/or viability may be altered. Markers to detect change in these aspects of cell behaviour were therefore employed.

Analysis of cell proliferation and programmed cell death

Cell proliferation was assayed using an antibody that detects cells in metaphase. The antibody used detects phosphorylated histone H3 (H3P) and begins binding weakly to pericentromeric heterochromatin of cells in G2 but the signal spreads as chromosomes condense becoming strongest during metaphase (Hendzel et al., 1997). Throughout brain development this signal is normally restricted to dividing cells at the ventricular surface but its distribution is altered where mutation interferes with cellular dynamics (Estivill-Torrus et al., 2002). The number and distribution of H3P positive cells was examined at 11.5dpc, 14.5dpc and 16.5dpc in $Sox1^{m1-/-}$ /$Sox2^{\beta geo-/+}$ brain sections and littermate controls. No difference either in the number of metaphase cells marked or their distribution could be detected between these samples (data not shown).

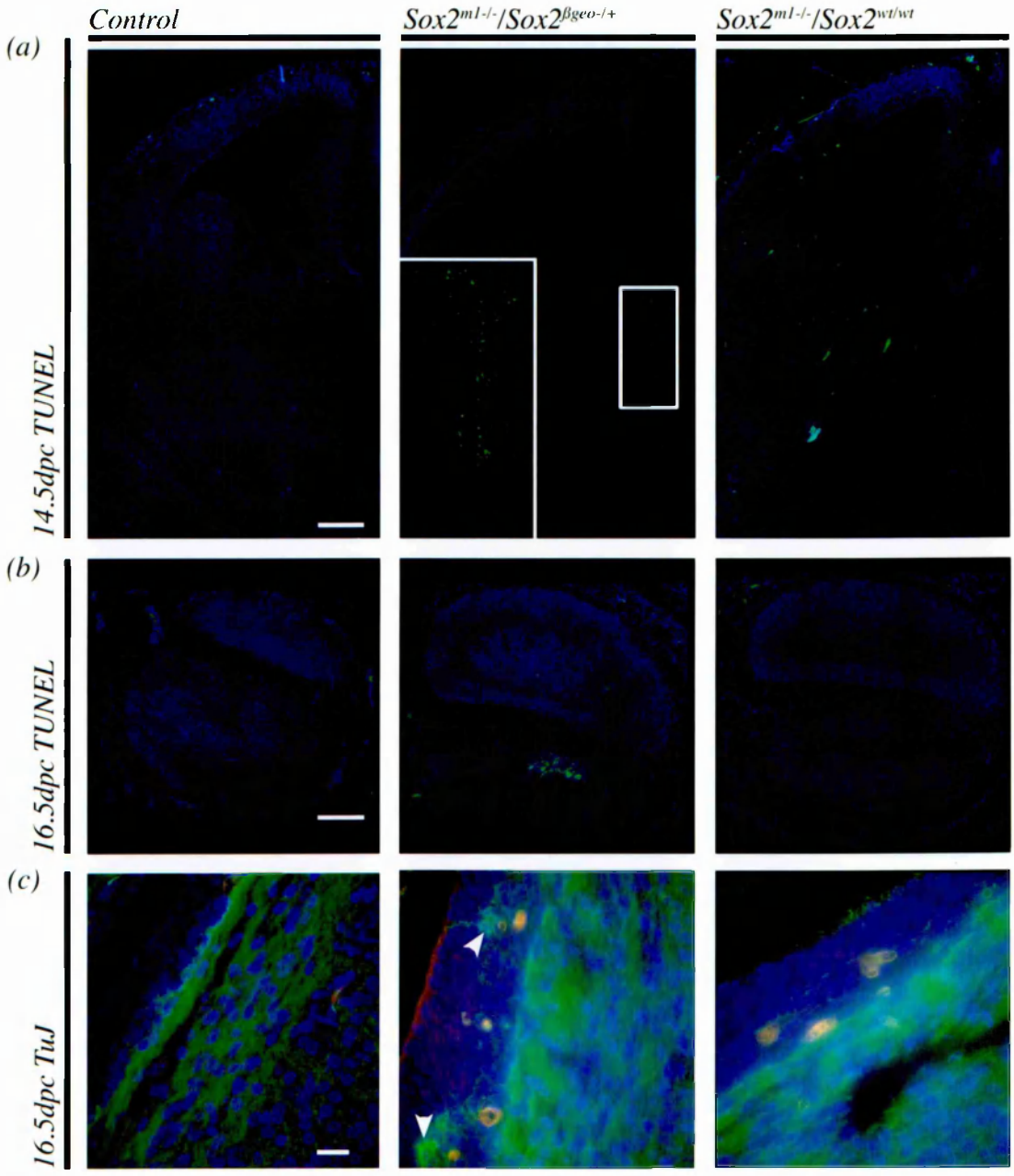
TUNEL analysis.

Another reason for a reduced number of SOXB1 positive cells might be that the null alleles cause cellular changes that compromise their viability. In order to detect cells that were undergoing programmed cell death TUNEL (terminal deoxynucleotidyl transferase dUTP nick end-labeling) analysis was carried out. This technique labels the ends of fragmented DNA that are produced within cells undergoing apoptosis (materials and methods).

Samples from the three different embryonic stages were examined (11.5dpc, 14.5dpc and 16.5dpc) and in general no differences in the number of apoptotic cells could detected. There was, however, some localised cell death that was only observed in the older $Sox1^{m1-/-}/Sox2^{\beta geo-/+}$ samples. In each case apoptosis occurred in

Figure 5.5 Unique differences are revealed in *Sox1$^{m1-/-}$/Sox2$^{\beta geo-/+}$* embryos by TUNEL and TuJ1 analysis.

(a),(b). *TUNEL analysis was performed on 14.5 and 16.5dpc Sox1$^{m1-/-}$/Sox2$^{\beta geo-/+}$ embryos alongside littermate controls. Clusters of TUNEL positive cells were observed in the undifferentiated ventricular zone where there is normally expression of the SoxB1 genes. Although the regions affected were not the same these clusters of cells marked by this procedure were only ever observed in Sox1$^{m1-/-}$/Sox2$^{\beta geo-/+}$ samples.* ***(a),*** *inset shows a close up of positive cells in the 14.5dpc forebrain in sagittal sections around the anterior margin of the third ventricle. Bar is 500μm.* ***(b),*** *TUNEL positive cells in the 16.5dpc Sox1$^{m1-/-}$/Sox2$^{\beta geo-/+}$ sample were located in the ventricular zone of the olfactory bulb. Bar is 500μm.* ***(c).*** *Preliminary immunohistochemical analysis using the TuJ1 antibody (to detect IIIß-tubulin in early differentiated neurons) marks ectopic expression in the ventricular zone of lateral ventricles only in Sox1$^{m1-/-}$/Sox2$^{\beta geo-/+}$ samples (arrowheads). Sections are in the region of the ganglionic eminence adjacent to the lateral ventricle of the forebrain. Bar is 20μm.*


Control
Sox2m1-/-/Sox2βgeo-/+
Sox2m1-/-/Sox2wt/wt
(a)
14.5dpc TUNEL
(b)
16.5dpc TUNEL
(c)
16.5dpc TuJ

the ventricular zone but the regions affected were not in the same areas. In one sample a group of apoptotic cells was detected adjacent to the ventricle of one of the olfactory bulbs (Fig. 5.3a). In a second sample olfactory neuroepithelium appeared normal but there was increased apoptosis in the anterior ventricular zone of the third ventricle. Such localised cell death may not result in gross morphological differences but could compromise the circuitry of the developing CNS and therefore the viability of animals carrying these lesions.

Preliminary analysis with cell type specific markers.

A preliminary survey was made using antibodies to mark specific cell types. Antibodies directed against radial glia (RC2), early migrating cells (DCX1) and early neurons (TuJ1) were employed to search for differences in these cell populations. In these initial experiments no differences attributable to the $Sox1^{m1-/-}/Sox2^{\beta geo-/+}$ phenotype were observed with the RC2 or DCX1 antibody, however, when using the TuJ1 antibody, which detects IIIß-tubulin, some alterations were seen. Preliminary analysis indicated that cells ectopically expressing this marker could be found within the ventricular zones only in $Sox1^{m1-/-}/Sox2^{\beta geo-/+}$ samples. This antibody was only tried once and so further experiments are necessary in order to confirm this result (Fig. 5.5c).

5.4 Discussion

Sox1, Sox2 and *Sox3* are grouped together by virtue of a high degree of sequence similarity but they also have extensive overlap in their expression patterns (Chapter 1). Animals null for each of the *SoxB1* genes indicate the importance of their protein products during development but phenotypes are generally limited to tissues where there is unique expression of a single member of this subgroup (Chapter1). The limited consequence of the null mutation of a single gene (at least for *Sox1* and *Sox3*) despite a broad expression pattern suggests functional compensation by related proteins with overlapping expression. If this is the case then more than one gene must be removed in order to decipher their role. Animals carrying various combinations of null alleles of *Sox1* and *Sox2* were therefore created to address the possibility of functional redundancy between these genes.

In this experiment all the expected genotypes were produced and were able to survive until birth but $Sox1^{m1-/-}/Sox2^{\beta geo-/+}$ pups always died before weaning age, most commonly soon after birth. The summation of the individual phenotypes observed in animals carrying equivalent mutations in either *Sox1* or *Sox2* can not explain the postnatal lethality of $Sox1^{m1-/-}/Sox2^{\beta geo-/+}$ animals and is an indication that there is synergy between these genes.

The expression of *Sox1, Sox2* and *Sox3* in the developing CNS indicates that they may have a role in maintaining progenitor cell identity before they start to migrate away from the ventricular zones. The overlapping expression of all three genes is highly conserved and so is likely to be a requirement for normal development. The introduction of null alleles may reduce the total complement of

these genes, but expression from the remaining wild-type *SoxB1* genes may broadly, but not completely, compensate for the loss.

Immunohistochemical analysis indicated an alteration in the extent of SOX2 and SOX3 expression from the remaining wild-type alleles in $Sox1^{m1-/-}/Sox2^{\beta geo-/+}$ samples. *SoxB1* proteins were also generally more difficult to detect as the number of null alleles present increased. This might reflect a degree of auto/crossregulation of these *Sox* genes, which has been previously suggested, exacerbating the effect of the null alleles present (Graham et al., 2003; Tomioka et al., 2002; Uchikawa et al., 1999; Wiebe et al., 2003). Assuming that SOX2 and SOX3 still mark cells of the same identity then the overall reduction in *SoxB1* expression has the effect of reducing the population size of this cell type. There could be a number of explanations for this such as a decrease in proliferation, an increase in cell death or the accelerated differentiation of *SoxB1* expressing cells. Initial analysis surveyed the forebrains of mutant embryos at various stages to identify any differences in cell proliferation or viability.

When the number and distribution of metaphase cells was examined there appeared to be no difference between samples. Also histological analysis did not indicate any obvious atrophy in comparable sections. This indicates that the stem cell population, present within the region marked by *SoxB1* expression (Pevny and Rao, 2003), is behaving normally and does not indicate that there is a generalised decrease in the rate of proliferation of progenitor cells.

TUNEL analysis was employed to identify cells undergoing programmed cell death but no general differences were observed between samples. The only

difference that could be attributed to the $Sox1^{m1-/-}/Sox2^{\beta geo-/+}$ genotype was the appearance of localised patches of TUNEL positive cells in the ventricular neuroepithelium of older embryonic samples.

Increased apoptosis is observed in a number of animal models when *Sox* genes are completely removed. Mice null for *Sox17* exhibit apoptosis of the endoderm in the foregut (Kanai-Azuma et al., 2002). Lack of any functional *Sox9* causes extensive programmed cell death in mutant limb buds although this is not true in animals harbouring a mixture of null and wild-type cells, indicating it is not an inherent property of deficient cells (Akiyama et al., 2002; Bi et al., 1999).

Increased apoptosis is also seen when the *SoxE* protein SOX10 is absent from neural crest cells, but haploinsufficiency leads to a fate change without increased apoptosis (Cheung and Briscoe, 2003; Paratore et al., 2001). The *SoxE* relative SOX9 is also found in neural crest and in *Xenopus* appears to be required there for the initial expression of *Sox10.* Later in development these two genes resolve a mutually exclusive pattern of expression and cranial neural crest development is upset, in a fashion similar to the null *Sox10,* when expression of *Sox9* is disrupted (Spokony et al., 2002). This suggests that some *SoxE* expression is required for neural crest cell survival but the correct amount is necessary for cells to make appropriate fate decisions. This is similar to the situation in animals carrying various combinations of null alleles of the identically expressed *SoxD* genes *Sox5* and *Sox6* (Smits and Lefebvre, 2003). Increasingly severe abnormalities in notochord development are observed as more null alleles are introduced but widespread apoptosis is seen only when both genes are absent (Smits and Lefebvre, 2003). This suggests that although

there may be some unique activity, shared functions between these two proteins ensures that basic structural formation can be achieved as long as some protein is present. Refinement of this basic structure is, however, reliant upon the expression of the correct quantities of these genes. This quantitative regulation will have been achieved through evolution. Evolution has also resulted in some unique function of each of these proteins so that it is not just a matter of achieving the correct absolute amount but some of each may be required for optimum development even in tissues where there is overlapping expression.

In this experiment dying cells observed in the ventricular zones of $Sox1^{m1-/-}/Sox2^{\beta geo-/+}$ samples do not occur frequently enough to account for the more general reduction of *SoxB1* positive ventricular cells. This difference may be due to the reduction in the overall dose of these genes affecting cell dynamics in $Sox1^{m1-/-}/Sox2^{\beta geo-/+}$ CNS. A reduced dose, although usually sufficient for survival, might trigger premature differentiation and/or the adoption of inappropriate cell fates as seen in the *Sox10* knockout. A general effect such as this might be the cause of serious neuronal defects leading to postnatal lethality when *Sox1* and *Sox2* alleles are disrupted.

The mammalian *hairy* and *enhancer of split* homologue-1 (*Hes1*) is a basic helix-loop-helix (bHLH) transcription factor that is expressed at high levels in the ventricular zone of the developing CNS. Homozygous disruption of *Hes1* causes the increased lethality of embryos older than 12.5dpc and no animals survive more than one day postpartum. Lethality is attributed to the premature and disproportionate differentiation of progenitors to a neuronal fate as well as increased apoptosis of

neural precursors in the ventricular zones (Ishibashi et al., 1995; Nakamura et al., 2000). Null mutation of *Hes5* is not lethal under normal laboratory conditions and again the morphological arrangement of the forebrain appears normal (Ishibashi et al., 1995; Ohtsuka et al., 1999). Most *Hes5* function appears to be compensated for, but similar accelerated differentiation is seen in the ventricular zone of the *Hes5* null CNS. A degree of functional compensation is revealed when compound mutants are made lacking both *Hes1* and *Hes5* (Ohtsuka et al., 1999). The brains of these embryos are smaller and severely deformed; neurons appear at an even greater density but there are still many *Nestin* positive neural precursor cells present. This indicates that the initial stages of neurogenesis are able to proceed but that lack of *Hes* genes later disrupts the correct timing and differentiation of progenitors.

$Sox1^{m1-/-}/Sox2^{\beta geo-/+}$ embryos continue to have expression from one wild-type *Sox2* and one *Sox3* allele (the latter due to X-inactivation. Like the individual *Hes* mutants expression from related wild-type alleles may be sufficient to maintain basic, although not perfect, integrity of the regulatory network and therefore CNS structure. In wild-type animals background differences would not cause expression levels to fall below a threshold critical for viability, but in mutant embryos a deficit may occur in certain zones. In $Sox1^{m1-/-}/Sox2^{\beta geo-/+}$ mutants the pockets of apoptotic cells may therefore be the result of localised background dependant differences in already depleted wild-type *SoxB1* gene expression. It is possible that the further reduction of *SoxB1* levels could lead to more significant and reproducible programmed cell death, but alterations in the behaviour of $Sox1^{m1-/-}/Sox2^{\beta geo-/+}$ cells may yield greater insight into role of these proteins. The elimination of background differences by utilising samples from inbred mouse lines may yeild more consistant and perhaps extensive

areas of apoptosis. In these samples it may be easier to identify exactly which cells in the *Sox1*$^{m1-/-}$/*Sox2*$^{\beta geo-/+}$ CNS are most sensitive to the reduction in *SoxB1* expression.

Sox1$^{m1-/-}$/*Sox2*$^{\beta geo-/+}$ embryos display some similar phenotypes to the *Hes1* and *Hes5* knockouts and could potentially be involved in a similar aspects of fate determination. In order to test this possibility markers of downstream events in the differentiation pathway should be examined and might indicate accelerated differentiation. Subventricular zone markers such as *Dlx5* (Anderson et al., 1997) or *Svet1* (Thomas and Capecchi, 1987) may be useful to detect changes in the arrangement of these layers. If reduction of *SoxB1* genes leads to premature differentiation of progenitors then these markers may well show an expanded domain of expression. As cell fate may also change, cell type specific markers may show a shift in the distribution and proportion of cells arising from ventricular progenitors such as early neuronal (Neuron-specific class III b-tubulin TuJ1 or microtubule associated protein-2 MAP2) or glial (GFAP or GLAST (astrocytes) radial glia (RC2)).

Preliminary analysis with the early neuron specific antibody TuJ1 indicated that there may be premature differentiation of progenitor cells in the striatal ventricular zones of 16.5dpc *Sox1*$^{m1-/-}$/*Sox2*$^{\beta geo-/+}$ samples. These experiments were preliminary as they were only carried out on a single set of samples but they do indicate an interesting observation that requires confirmation. In chick neural tube the expression of a dominant negative form of *Sox2* also has the effect of accelerating the rate of differentiation of neuronal precursors, whilst over expression prevents differentiation without stimulating over proliferation (Graham et al., 2003).

The redistribution of early neuronal cells and increased cell death are phenotypes remarkably similar to that seen in samples null for either of the bHLH genes *Hes1* or *Hes5*. This indicates a role for each of these genes around the same stage of neurogenesis. It is possible that one influences the expression of the other or that each is required for the correct regulation of a set of common targets.

To date the only common target shown to be regulated by *Sox* and bHLH factors is the fruit fly gene *Slit.* The function of this family of extracellular matrix molecules is evolutionarily conserved with repulsive activity necessary for appropriate axon guidance and possibly neuronal migration (Plump et al., 2002). In *Drosophila, Slit* gene regulation requires the concerted contribution of the bHLH-PAS gene *Singleminded* (*Sim*), the POU domain protein *Drifter* (*Dfr*) and the *Sox* gene *Dichaete* (*Fish, Fish-hook*, SOX70D, SOXB2.1) interacting directly on the *Slit* enhancer (Ma et al., 2000). *Sim* is an absolute requirement for the initiation of *Slit* expression whilst its interaction with *Sox*/POU complexes is necessary for maintenance of *Slit* expression (Ma et al., 2000). Interaction in this case is between the PAS (*Per, Arnt, Sim* (Crews and Fan, 1999)) domain of *Sim* and the HMG box of *Dichaete* but differential regulation may be achieved in the presence of alternative *Sox*, POU or bHLH proteins. It is possible that a similar regulatory relationship has been conserved as *Slit1* and *Slit2* have overlapping expression with all three classes of proteins in the ventricular zones of the developing mouse forebrain (Marin et al., 2002). It would therefore be interesting to examine the expression of *Slit* genes in embryos carrying compound mutations in *Sox1* and *Sox2.*

The lack of obvious phenotypes in animals carrying null mutations has often been explained as due to 'functional redundancy' between closely related proteins. Redundancy is defined as 'the state or fact of not being or no longer being needed' that in this context suggests an unnecessary excess of a particular activity. The lethality of mice carrying compound null mutations might be taken to reflect redundancy between these related genes, but experiments *in vivo* and *in vitro* indicate that some function may not be compensated for even by very similar proteins.

The subgroup members of the vertebrate Sox family are thought to have arisen by duplication of ancestral genes. Initially, pairs of identical genes would rapidly accumulate mutations (within a few million years) refining activity and/or expression levels close to what is actually required. These alleles become fixed in the population when further mutation is detrimental to the host. The *SoxB1* proteins are extremely well conserved even between species where the closest ancestor occurred nearly two hundred million years ago and some functional conservation can even be demonstrated between mouse and fly Sox genes (Chapter1), (Ma et al., 2000; Soriano and Russell, 1998). Differences between these genes may in certain contexts confer different activities and the combination of common and unique activities may be the strength behind their contribution to the regulatory network. This combination will be a function both of expression pattern and protein activity and so a full complement is likely to be critical for optimal fitness. This is perhaps why even a single null *Sox* allele can cause developmental changes despite overlapping expression of a related protein (Smits and Lefebvre, 2003; Sock et al., 2001). Close examination may reveal the contribution of a single allele but minor developmental changes may have no

consequence to the behaviour or viability of the mollycoddled laboratory mouse and so might appear, on the face of it, insignificant or indeed redundant.

Several null mutations may be required before any obvious effect occurs in a tissue of interest, not due to redundancy per se, but due to the reduced relative contribution of a single allele. For example in the ventricular zones, *Sox1, Sox2* and *Sox3* are expressed giving a total of 5 expressed alleles (*Sox3* is known to be subject to X-inactivation (Rizzoti et al., 2004). Assuming that function and expression from each allele is equivalent then $Sox1^{ml-/-}$ would only remove two fifths of the necessary protein in sites where there is overlapping expression. Introduction of further null alleles, as in $Sox1^{ml-/-}/Sox2^{\beta geo-/+}$ mutants, would remove slightly more than half. This is equivalent to haploinsufficiency (or worse) that is capable of causing phenotypes such as increased postnatal lethality or hair follicle dysgenesis in $Sox2^{\beta geo-/+}$ animals. As mentioned above this effect may be exacerbated further if the correct levels of *Sox* genes are also required for auto or cross regulation.

Although a phenotype was observed in cells where there is overlapping expression of *Sox1, Sox2* and *Sox3* it is not possible to conclude that this is the cause of postnatal lethality of $Sox1^{ml-/-}/Sox2^{\beta geo-/+}$ animals. The cause of death may actually be due to defects in other parts of the CNS or quite separate phenotypes that synergistically result in the death of animals in this experiment.

The creation of a conditionally null *Sox2* allele and using it in the place of $Sox2^{\beta geo-/+}$ would allow a more refined reduction or elimination of *SoxB1* genes. Restricting gene disruption to certain parts of the CNS would make it easier to analyse the role of these genes in cells where there is overlapping expression. The

further introduction of the conditionally null allele of *Sox3* may make it possible to completely ablate the normal expression of all three *SoxB1* genes. It would also provide the opportunity to examine the relative importance of these proteins by creating animals lacking different combinations of *SoxB1* genes.

Whilst the cumulative importance of this subgroup of genes may be examined in this way it is unable to provide information about the unique functions of these genes that have evolved. In order to examine the role that functional compensation might play between close relatives, such as *Sox1, Sox2* and *Sox3, in vivo* the coding region of one gene may be swapped for another by gene targeting. In this case total dose with respect to equivalent function would remain unchanged and unique roles would be exposed as the replacement gene would be unable to compensate.

A mouse model attempting to answer this question has already been produced where the coding sequence of *Sox2* replaces its closest relative *Sox3* ($Sox3^{Sox2}$). Offspring display none of the severe deformities attributed to a null *Sox3* allele demonstrating major, although perhaps not complete, functional compensation by SOX2 (Silvia Brunelli unpublished data), (Parsons, 1997). A second targeted replacement where *Sox2* coding sequence is exchanged by that of *Sox1* ($Sox2^{Sox1}$) has also been created (Vasso Episkopou unpublished data). This model is homozygous lethal at the same stage as $Sox2^{\beta geo-/-}$ but the removal of critical regulatory sequence, preventing proper expression of the replacement gene early in development, negates drawing any conclusions about potential functional compensation. However, this model may still be useful as part of allelic series with the *Sox2* regulatory mutants *Ysb* and *Lcc* (Chapter 4). By creating animals carrying one each of $Sox2^{Sox1}$ and *Lcc*

or *Ysb* the degree of functional compensation by *Sox1* may be examined in the development of the inner ear and the hair follicle. A range of phenotypes might be observed depending on the level of functional compensation *Sox1* is able to provide in the place of *Sox2* and opens up screening possibilities to discover differentially expressed targets.

Chapter 6 Final Discussion

One of the most interesting aspects of *Sox2* is its relationship to other members both within its subgroup and to the rest of the family. Its closest relatives *Sox1* and *Sox3* share considerable homology throughout their sequence suggesting very similar functional characteristics. Experimental evidence reinforces this idea as well as revealing considerable overlapping expression. This has led to the suggestion that the overall dosage of these proteins is more important than any distinct function. Members of the other Sox subgroups share similarity within the HMG box with other domains being more diverse. Since target specification is fundamentally dependant upon DNA binding via the HMG domain one could imagine that this would allow any SOX protein to regulate the same target. Whilst similarities might permit certain targets to be regulated in a comparable way differences provide the means to regulate other targets differentially. This could result in certain SOX proteins having antagonistic effects.

SOX1, SOX2 and SOX3 are unequivocally very similar physically, functionally and in terms of expression pattern. The results presented in chapter five demonstrate that in regions where there is overlapping expression of the *SoxB1* genes obvious phenotypes are only revealed by the disruption of multiple subgroup members. A fundamental role in the ventricular zone was uncovered by the creation of $Sox1^{ml-/-}/Sox2^{\beta geo-/+}$ mutants. Whilst the gross morphology of the CNS in these samples appeared normal, immunohistochemistry revealed changes in cell viability and perhaps premature differentiation. These data indicate a basic role in the dynamics of the progenitor cell population and further analysis of these mutants may

help to explain the mechanisms more precisely. Initial investigations should focus on describing further characteristics of cells in the mutant ventricles especially with respect to differentiation markers. A more comprehensive survey of changes in the expression profile could be made by utilising microarray technology. It is possible that further phenotypes in *Sox1*$^{m1-/-}$/*Sox2*$^{\beta geo-/+}$ samples as well as those in other genotypes are masked to some extent because all the cells carry the same mutation. A more reliable (and comparable) assay of any functional differences might be achieved by examining mutant cells in the context of a wild-type environment. This could be accomplished *in vitro* by culturing wild-type embryonic brain slices onto which labelled experimental cells could be placed and monitored. The experimental cells would include wild-type as well as cells carrying combinations of all the available mutant alleles. Each brain slice could accommodate differently labelled cells harvested from the ventricular zone that would then be monitored in real time. This would allow the direct comparison of cell behaviour between wild-type and mutant samples in the same environment and would permit the analysis of mutant cells beyond that achievable *in vivo*.

Although *Sox1*$^{m1-/-}$/*Sox2*$^{\beta geo-/+}$ embryos provide useful material to study the role of *SoxB1* in the ventricular zone, extra information could be gained by disrupting further alleles. The conditional disruption of *Sox2* would allow its complete removal in the ventricular zone and similarly with the *Sox3* conditional knockout. The advantage of a conditional strategy is that one is not observing secondary effects. These experiments could be carried out *in vivo* by the unilateral introduction of *cre* into the ventricular layer by *in utero* electroporation. By adopting this method *Sox1*, *Sox2* and *Sox3* could all be removed in a localised cell population. This is likely to

produce the most severe phenotype and give a definite insight into the role of *SoxB1* genes/proteins in sites of overlapping expression.

Currently there is evidence that the activity of *Sox1, Sox2* and *Sox3* is very similar, however, equivalent function is not consistent with evolutionary selection pressures. The conservation of overlapping patterns of expression and variations in protein structure over a long period of evolutionary time strongly supports a selective advantage in maintaining subtle differences. Transcription factors with similar but not identical activities offer a sophisticated strategy to refine the expression characteristics of a cell. Therefore it is important to distinguish between functional equivalence and functional similarity. Conventional approaches make it very tricky to differentiate between these possibilities because of the difficulty in speculatively screening large numbers of potential targets for subtle differences. Microarray technology, however, offers a method by which the expression of thousands of genes may be analysed simultaneously and provides a quantifiable method for the comparison of different samples. Expression analysis demonstrates the overlapping expression of *SoxB1* genes in the ventricular zone of the developing CNS that is conserved between mouse and chicken. Microarray analysis of the expression profile of these cells using mutants heterozygous for *Sox1, Sox2* and *Sox3* and comparing to wild-type may reveal both the targets of these genes as well as subtle differences in their activity. It is precisely these subtle differences that are likely to have permitted increased complexity in the developing vertebrate CNS.

Whereas the similarity between *SoxB1* genes is likely to contribute to a subtle refinement of the regulatory network more dramatic differences may be brought

about by members of other subgroups. One advantage of conserving the core DNA binding domain across a large family of genes is that it allows diverse regulatory possibilities. Even in the most primitive animals there is a representative of each *Sox* subgroup and these have varying degrees of overlapping expression. This prompts the suggestion that certain regulatory processes might critically rely upon the antagonistic behaviour of related proteins and that these fundamental relationships have been crucial over a long period of time.

The mouse mutants *Lcc* and *Ysb* were used along with $Sox2^{\beta geo}$ to demonstrate the requirement of *Sox2* for the correct development of the inner ear and the hair follicle. In the hair follicle relatively mild phenotypes were seen but what is most striking is the similarity of phenotypes observed after the null mutation of the group F gene *Sox18*. Hair type and pigmentation characteristics were altered in both instances but the effects in each case were opposite to each other. This is intriguing because it suggests that the appearance of the mouse coat is dependent upon a balance between these related factors and it is possible that they may regulate some common targets.

Furthermore dominant negative forms of *Sox18* expressed by *ragged* alleles cause far more severe coat abnormalities revealing significant roles for *Sox* factors in hair follicle development. So far apart from *Sox2* and *Sox18* no other *Sox* gene expression has been described during hair follicle development and so the *ragged* phenotype could be solely due to a disruption of their regulatory responsibilities. It would therefore be interesting to reduce the amount of both these genes in the same animal by breeding together the null mutants to see if a *ragged*-like phenotype can be reproduced. If the coat phenotype of *Sox2/Sox18* null animals is not the same then

further *Sox* genes, susceptible to interference by *ragged* protein, are likely to be responsible for vital aspects of follicle development. However, the replication of a *ragged*-like phenotype confirms that the correct expression of these genes is critically important but also suggests that whilst they regulate some targets antagonistically others may respond in a similar way. An insight into the molecular mechanisms underlying these processes may be revealed by comparison of follicular material using the multi-gene screening techniques discussed above. It is even possible that this could represent a common regulatory strategy in many of the sites where there is overlapping expression of related transcription factors. If this is the case then the hair follicle may be used as a convenient model system in which to examine this phenomenon. Phenotypes arising from genetic manipulations are readily assayed and severe disruptions, unlike those in the CNS, are unlikely to compromise viability. By exploiting this system the mechanisms of regulation by the SOX proteins could be studied *in vivo* allowing functional differences between them to be examined. For example the over-expression of *Sox2* in the dermal papilla could lead to phenotypes similar to those of the *Sox18* null by altering the balance between their two proteins. The expression of other members of the *Sox* family may reveal how their activity compares in the same context and could go some way towards explaining the regulatory consequences of expressing different combinations of these transcription factors *in vivo*.

Analysis of further mutant alleles is likely to be of key importance in understanding the kind of regulatory interactions described above. The creation of conditionally null mutants provides the means with which to analyse function whilst avoiding many of the pitfalls associated with traditional gene disruptions such as

early lethality and secondary phenotypes. However, although general roles are revealed by the complete disruption of specific alleles the dissection of an individual gene's activity in an *in vivo* context is likely to require the isolation or production of even more subtle mutations. For example the targeted replacement of one subgroup member by another may reveal the how slight physical differences contribute to optimal function, especially in sites where there is normally overlapping expression.

Sox2 is clearly crucial for many aspects during mouse embryogenesis as well as postnatal development. This thesis aimed to clarify the regions of *Sox2* expression in the second half of gestation and highlights novel sites including the hair follicle and inner ear. These areas have no overlapping expression of other *SoxB1* genes and this enabled the study of SOX2 function alone. The investigation of compound mutants revealed phenotypes that suggest the SOXB1 proteins work together during neurogenesis in the CNS. It is therefore important not to consider each one alone whilst still being aware of the potential for distinct functions. The advent of new technology along with the generation of more sophisticated mouse mutants will aid the unravelling of these complex regulatory interactions.

Chapter 7 Materials and Methods

Standard molecular biology techniques were carried out essentially as described in Molecular Cloning, A Laboratory Manual and Current Protocols in Molecular Biology (Ausubel, 1999; Sambrook, 1989) and reagents used as per manufacturers recommendations unless otherwise stated.

7.2 Screening of transgenic mice

Standard PCR was performed using Perkin Elmer *amplitaq* as per the manufacturers recommendations. Long PCR was performed using the Boehringer Mannheim *Expand* long PCR kit.

7.2.1 PCR analysis of *Ysb* alleles.

The detection of *Ysb* alleles was carried out as detailed in Dong et.al., 2002.

7.2.2 Maintenance and screening of *Lcc* mouse lines.

Lcc mouse lines were screened by test breeding each mouse. Heterozygote animals do not have circling behaviour, but produce circling offspring at a normal Mendelian ratio when bred together, thus confirming parental genotype. A number of homozygote males (circling) proved to be fertile and were used, particularly in compound breeding experiments, to guarantee the presence of an *Lcc* allele.

7.2.3 PCR analysis of *Sox1 null* allele.

A cocktail of primers was used to detect the null and wild type allele of *Sox1*. NEO2 (5' CTT CCT CGT GCT TTA CGG TAT CGC 3') (mutant allele) SX13'R (5'TGA TGC ATT TTG GGG GTA TCT CTC 3') (wild type allele) and SX3Fnew (5'TTA CTT CCC GCC AGC TCT TC 3')

(common antisense oligo) gave bands of 480bp for the null allele and 373bp for the wild type allele.

7.2.4 PCR analysis of *Sox2 null* allele.

The primer pair SOX2 2BS (5'CAC AGT CCT GGC CGG GCC GAG G 3') and LACZPZ4A (GTA GAT GGG CGC ATC GTA ACC GTG C) was used to detect the null allele of *Sox2* and gives a band of 336bp. Screening of cells targeted with pDB13 was carried out using the above primer set and then oligos ODB25 (5'CAT CGC CTT CTA TCG CCT TCT TGA 3') and ODB28 (5'CAA GCC CCT TCC TTG TTC CT 3') were used for long PCR analysis. Control reactions were performed using $Sox2^{\beta geo+/-}$, 129/Sv genomic DNA and no DNA.

7.3 Method for taking and examining hair samples

Three hair samples were taken from each mouse in the mid dorsal region just posterior to the shoulder blades. Each hair pluck was carried out in a single movement using a pair of Millipore hybridisation forceps (smooth). All samples were then counted blind before collation.

7.4 Targeting and screening of Sox2 in ES cells.

7.4.1 Construction of the vectors used in targeting experiments.

The *Sox2* conditional knockout targeting vectors were constructed using genomic DNA from a male 129SvEv mouse lambda library.

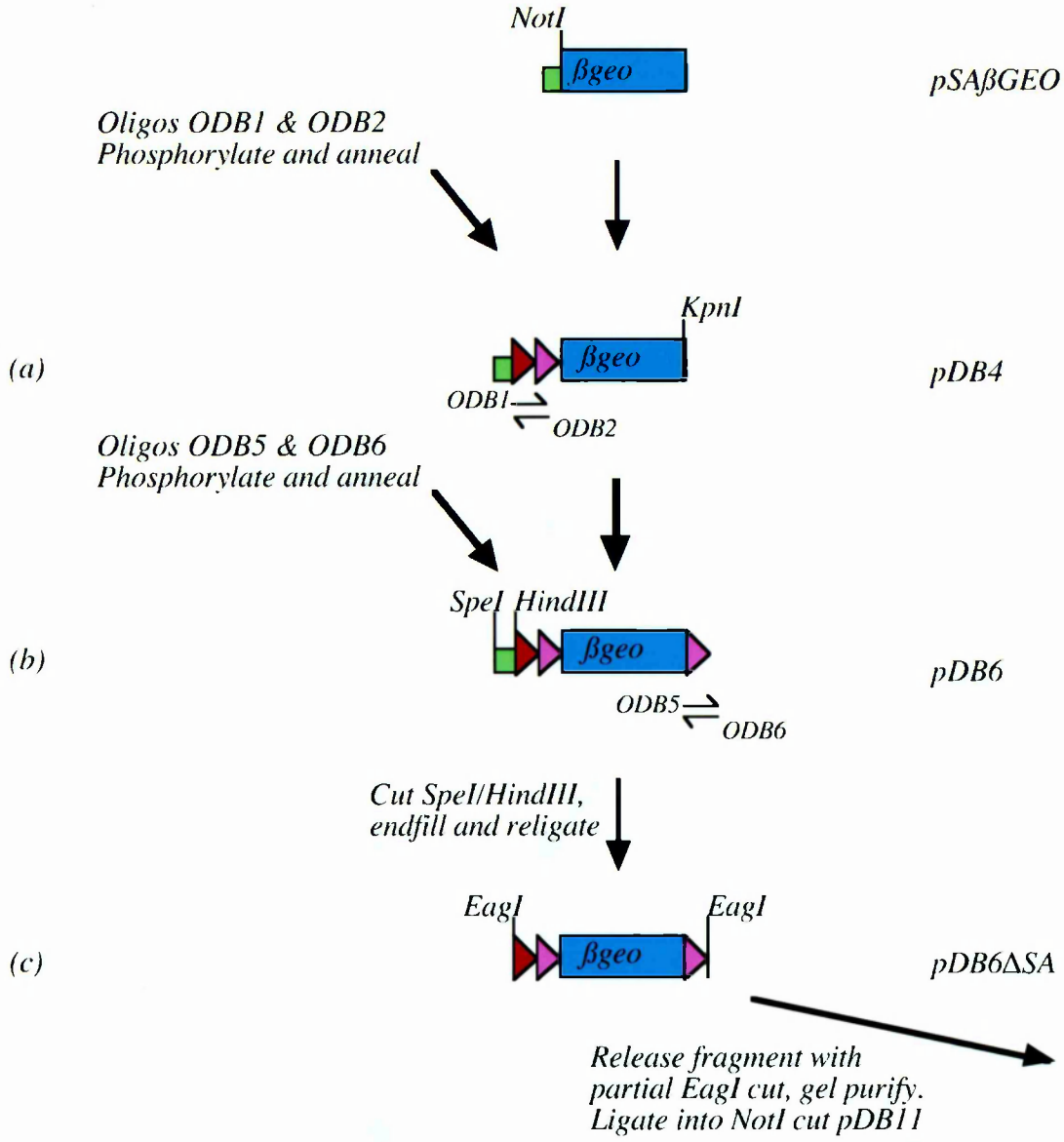
7.4.2 Phosphorylation and annealing of complementary oligonucleotide pairs used in cloning.

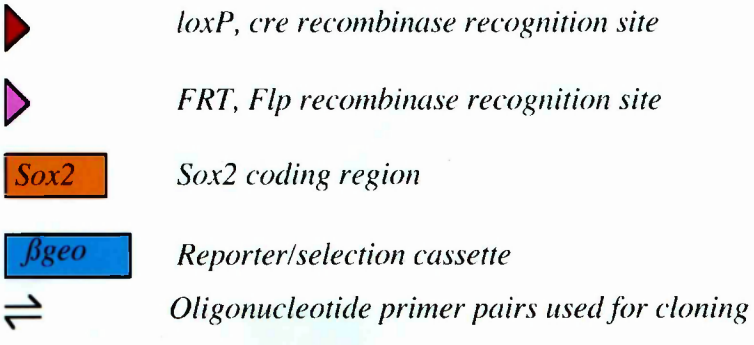
All oligonucleotides were synthesised and HPLC purified by MWG Biotech. Pairs used for cloning were diluted in ligase buffer (New England Biolabs) and phosphorylated by the addition of T4 polynucleotide kinase (PNK (New England Biolabs)). PNK was then inactivated by incubating at 75°C for ten minutes and annealing of the oligonucleotides was then carried out directly by raising the temperature to 95°C for one minute then allowing the reaction to cool to room temperature. Where pairs were to be ligated an approximately equal molar quantity of phosphorylated and annealed oligonucleotide pair and vector was used in the reaction.

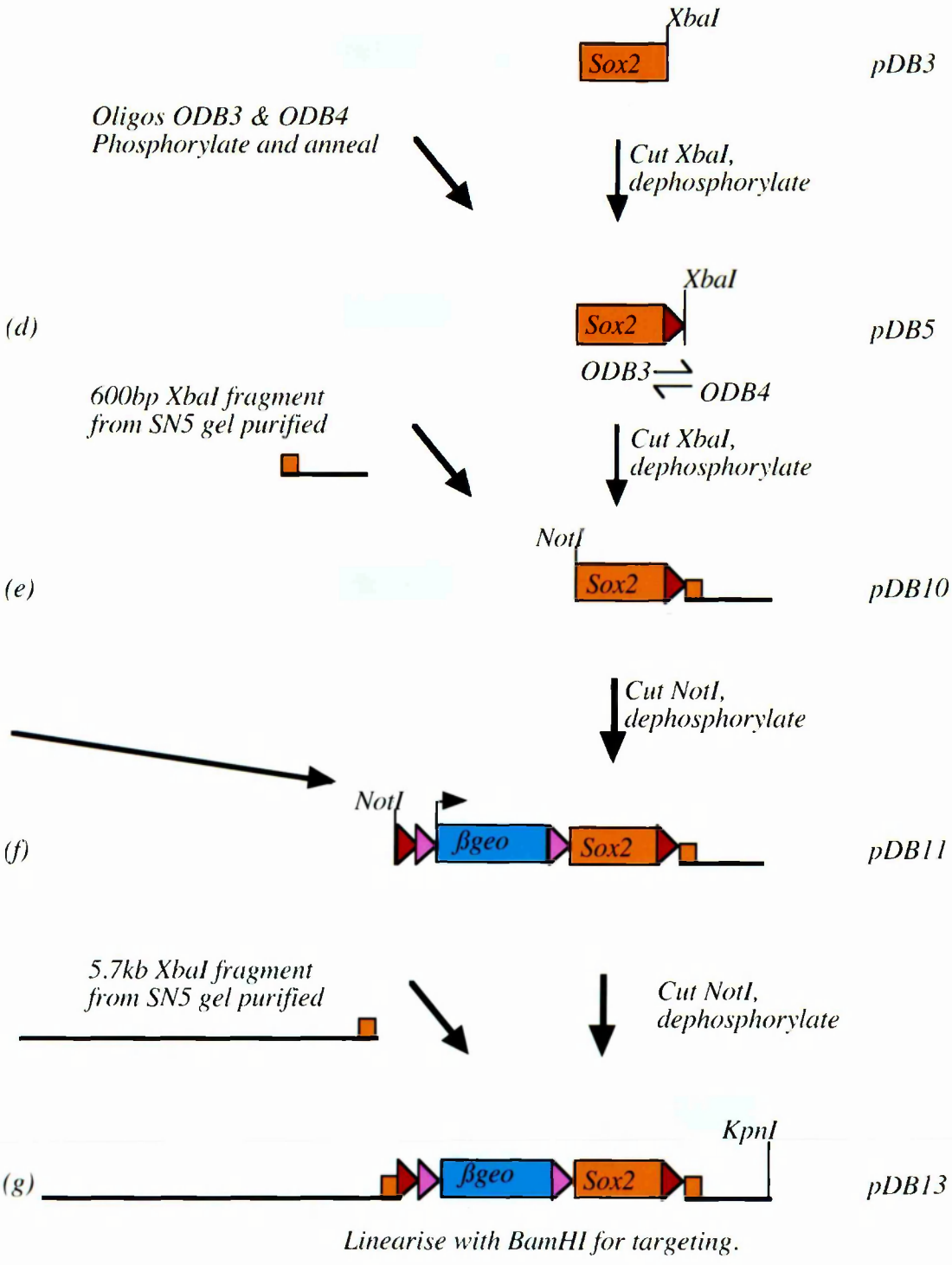
7.4.3 Production of the first targeting vector pDB13.

The oligonucleotide pair ODB1 (5'GGC CGC ATA ACT TCG TAT AGC ATA CAT TAT ACG AAG TTA TGA AGT TCC TAT TCC GAA GTT CCT ATT CTC TAG AAA GTA TAG GAA CTT C 3') and ODB2 (5'GGC CGA AGT TCC TAT ACT TTC TAG AGA ATA GGA ACT TCG GAA TAG GAA CTT CAT AAC TTC GTA TAA TGT ATG CTA TAC GAA GTT ATG C 3') was ligated with the plasmid that contained the beta-galactosidase, neomycin fusion gene pSAßGEO (Friedrich, 1991) which had been cut with NotI and then dephosphorylated with calf alkaline phosphatase(CIP (Roche)). This step introduced the cre and flp recombinase recognition sites, *loxP* (Smith, 1995) and *frt* (O'Gorman, 1991) 5' of the selection cassette producing pDB4 (Fig. 7.1a). A single *frt* site was then introduced, using the oligonucleotide pair ODB5 (5'GAA GTT CCT ATT CCG AAG TTC CTA TTC TCT AGA AAG TAT AGG AAC TTC GGC CG 3') and ODB6 (5'CGG CCG AAG TTC CTA TAC TTT CTA GAG AAT AGG AAC TTC GGA ATA GGA ACT TCG TAC 3'). This

Figure 7.1 Production of the first targeting construct pDB13.


NotI
βgeo
pSAβGEO
Oligos ODB1 & ODB2
Phosphorylate and anneal
KpnI
(a)
βgeo
pDB4
ODB1
ODB2
Oligos ODB5 & ODB6
Phosphorylate and anneal
SpeI HindIII
(b)
βgeo
pDB6
ODB5
ODB6
Cut SpeI/HindIII,
endfill and religate
EagI
EagI
(c)
βgeo
pDB6ΔSA
Release fragment with
partial EagI cut, gel purify.
Ligate into NotI cut pDB11


loxP, cre recombinase recognition site
FRT, Flp recombinase recognition site
Sox2
Sox2 coding region
βgeo
Reporter/selection cassette
Oligonucleotide primer pairs used for cloning


XbaI
Sox2
pDB3
Oligos ODB3 & ODB4
Phosphorylate and anneal
Cut XbaI,
dephosphorylate
(d)
XbaI
Sox2
pDB5
ODB3
ODB4
600bp XbaI fragment
from SN5 gel purified
Cut XbaI,
dephosphorylate
NotI
(e)
Sox2
pDB10
Cut NotI,
dephosphorylate
NotI
(f)
βgeo
Sox2
pDB11
5.7kb XbaI fragment
from SN5 gel purified
Cut NotI,
dephosphorylate
KpnI
(g)
βgeo
Sox2
pDB13
Linearise with BamHI for targeting.

oligonucleotide pair was ligated into pDB4 that had been cut with KpnI and dephosphorylated with CIP to produce the plasmid pDB6 (Fig. 7.1b). The adenovirus splice acceptor site was then removed by cutting pDB6 with SpeI/HindIII, endfilling with DNA polymerase I (large Klenow fragment) followed by religation to produce the plasmid pDB6ΔSA (Fig. 7.1c).

pDB3 was cut with XbaI, and dephosphorylated using CIP. The oligonucleotide pair ODB3 (5'CTA GCA TAA CTT CGT ATA GCA TAC ATT ATA CGA AGT TAT 3') and ODB4 (5'CTA GCC TAG GAT AAC TTC GTA TAA TGT ATG CTA TAC GAA GTT ATG 3') was ligated into the pre-prepared pDB3 vector to produce the plasmid pDB5 (Fig. 7.1d). This step introduced a single *loxP* site 3' of the coding region of *Sox2*. A 600bp piece of *Sox2* genomic DNA was isolated by gel purification of the clone SN5 that had been cut with XbaI. This fragment was then ligated into pDB5 that had been cut with XbaI and dephosphorylated with CIP to produce pDB10 (Fig. 7.1e). pDB10 was cut with NotI and dephosphorylated. The *loxP/frt* flanked selection cassette was released from pDB6ΔSA by a partial EagI digestion that produces sticky ends compatible with NotI. Gel purification produced a 3.9 kilobase (kb) fragment that was then ligated into the NotI cut pDB10 vector to produce plasmid pDB11 (Fig. 7.1f). pDB11 was cut with NotI and a 5.7kb genomic DNA fragment from SN5 was introduced 5' to the coding region of *Sox2*, to produce the final targeting construct pDB13 (Fig. 7.1g). For electroporation into ES cells pDB13 was linearised with KpnI.

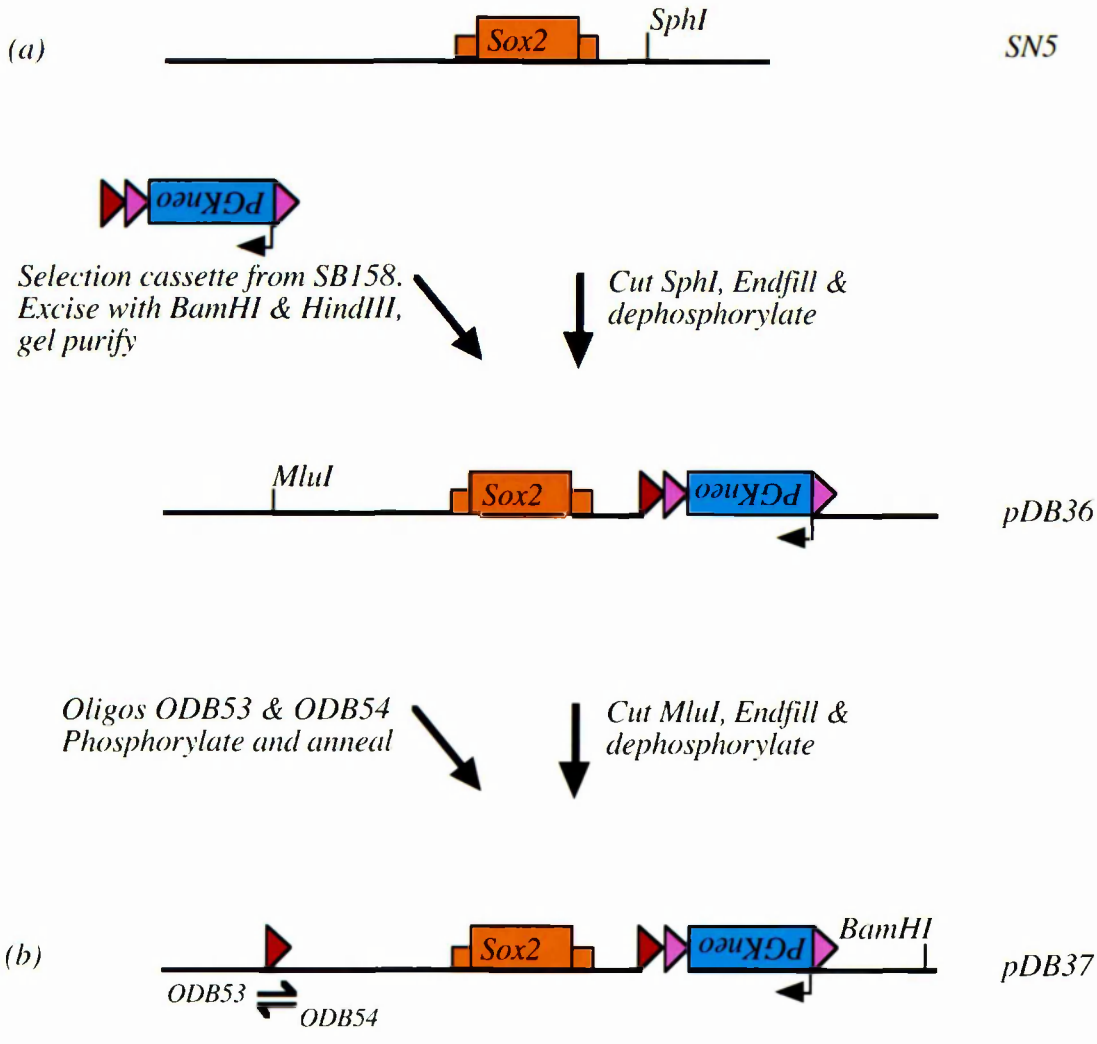
7.4.4 Production of the second targeting vector (pDB37)

SB158 was digested with the restriction enzymes BamHI and HindIII. Digested DNA was end filled with T4 DNA polymerase (Roche) and the 2kb

fragment, containing the PGKneo selection cassette, flanked by frt sites and followed by a *loxP* site, was gel purified. This fragment was then ligated into SN5 plasmid DNA that had previously been linearised with SphI, endfilled with T4 DNA polymerase (Roche) and dephosphorylated with CIP (Roche) producing pDB36 (Fig. 7.2a,b).

To introduce the second *loxP* site into the targeting vector a pair of complementary oligonucleotides was designed containing the *loxP* sequence flanked by MluI sites (ODB53 (5'CGC GTA CTA GTA TAA CTT CGT ATA GCA TAC ATT ATA CGA AGT TAT3') ODB54 (5'CGC GAT AAC TTC GTA TAA TGT ATG CTA TAC GAA GTT ATA CTA GTA3'). Oligonucleotides were ligated into pDB36 that had been linearised with MluI and dephosphorylated with CIP (Roche) (Fig. 7.2c). Promising clones of the completed construct were sequenced to confirm the presence of a single *loxP* site in the correct orientation. pDB37 was linearised using BamHI (Fig. 7.2c) before electroporating into ES cells.

Figure 7.2 Production of the second targeting construct pDB37



loxP, cre recombinase recognition site

FRT, Flp recombinase recognition site

Sox2 — Sox2 coding region

PGKneo — Reporter/selection cassette

Oligonucleotide primer pairs used for cloning

7.4.5 Culture of ES cells.

Embryonic stem cell lines used were either CCE (Bradley 1984) or AB1 (kind gift R Behringher). Both lines were derived from XY 129Sv/Ev pre-implantation mouse embryos. This mouse strain carries the agouti coat colour marker. ES cells were grown in Dulbecco's modified Eagles' medium (DMEM) (Sigma) supplemented with 2mM L-glutamine (Gibco), 100units/ml penicillin (Gibco), 100µg/ml streptomycin (Gibco), 90µM ß-mercaptoethanol (Sigma), 15% ES qualified foetal calf serum (FCS)(Globefarm), at 37°C in a humidified incubator with 5%CO_2. Media for AB1 ES cells also contained 1000 units/ml of leukaemia inhibitory factor (LIF) (Gibco). Cells were passaged by removing the media and washing in PBS (without calcium and magnesium) (PBS^{CMF})(Gibco). Just enough 1X Trypsin-EDTA solution (Gibco) was added to cover the cells, which were then incubated for five minutes at 37°C, until cells had lifted off the plate. An equal volume of media was added to the cell suspension, which was then centrifuged at 1000 rpm for five minutes to pellet the cells. The supernatant was aspirated off, cells resuspended in media and plated at the desired concentration.

ES cells were plated on top of a feeder layer of mitomycinC treated, neomycin resistant fibroblast cells (STO(N)) (Robertson 1987) to prevent differentiation. These were produced by incubating STO(N) cells in media supplemented with 1µg/ml mitomycinC (Sigma) for two hours. Cells were washed three times in PBS^{CMF}, trypsinised and plated onto gelatinised plates at a concentration of $3.8x10^4$ cells/cm^2. Media was changed daily on ES cell cultures.

7.4.6 Transfection of ES cells by electroporation.

Targeting vectors were linearised, phenol/chloroform extracted, ethanol precipitated and resuspended at a concentration of about 1μg/ml. ES cells were trypsinised and resuspended in PBS^{CMF} at a concentration of $3.33x10^7$ cells/ml. $2x10^7$ cells (0.6ml) were placed in a 0.4cm electrode Gene Pulser® curvette (BioRad) and electroporated at 960μF at 0.2 Volts. After electroporation cells were added to 10ml fresh media and split between five 10cm plates of STO(N) feeder cells. A control sample was processed in exactly the same way except no DNA was added.

7.4.7 Selection of transfected cells with antibiotic.

Electroporated ES cells were allowed to recover for two days, being maintained on plain ES cell media. On the third day after electroporation media was supplemented with 250μg/ml of the antibiotic G418 (Sigma). Selection media was used usually for about ten days and was regarded as sufficient when no ES cells remained on the control plates. By this time discrete, G418 resistant ES cell colonies were present on experimental plates and were of sufficient size to pick and culture separately. This was achieved by replacing media with PBS^{CMF} and picking individual colonies up with a Gilson pipette and sterile filter tips. Colonies were transferred, in a minimum volume, to a round bottomed, ninety-six well tissue culture plate that contained 50μl 1X Trypsin-EDTA (Gibco) in each well. Plates were transferred to 37°C for five minutes. 150μl fresh media was then added and the suspension was pipetted up and down several times to disperse cells. Cells were then transferred to a fresh ninety-six well plate and cultured. Once the majority of wells in a plate were confluent cells were passaged, one third of each to two ninety-six well

feeder plates for freezing and the final third divided onto two gelatinised ninety-six well plates to recover DNA for screening.

7.4.8 Freezing of ES cells in 96 well plates.

ES cell clones were frozen in ninety-six well plates, after aspirating the media and washing the cells with 200μl PBSCMF. 50μl 1X Trypsin-EDTA (Gibco) was added to each well and plates were incubated at 37°C for five minutes. 200μl freezing media (40% ESQ FCS, 20% dimethyl sulphoxide (DMSO)) was added to each well and cells were pipetted gently to ensure a single cell suspension. Each well was overlaid with 100μl light paraffin oil before sealing them with parafilm, wrapping in several layers of paper towels and placing at –80°C.

7.4.9 DNA isolation from ES cells

DNA was isolated from ES cells grown on gelatinised 96 well tissue culture plates. First 50μl of lysis buffer (10mM Tris pH7.5, 10mM EDTA pH 8.0, 10mM NaCl, 0.5% sarcosyl, 1mg/ml PK) was added to each well and incubated overnight at 60°C in a humidified polythene bag. DNA was then precipitated by adding 100μl, freshly prepared, 75mM NaCl in 100% ethanol and leaving at room temperature for thirty minutes. Liquid was discarded, by briskly inverting the plate and the remaining DNA precipitate was washed three times by repeated addition and discarding of 200μl of clean 70% ethanol. Plates were then air dried in an inverted position for thirty minutes before the addition of 25μl of 0.5TE (5mM Tris pH7.5, 1mM EDTA) to resuspend the DNA. To ensure proper resuspension of the DNA precipitate the plates were placed in a humidified bag at 37°C overnight.

7.4.10 Southern blotting

25μl 2X enzyme mix (2X buffer 2mM spermidine (Sigma), 200μg/ml BSA and 20-30U enzyme and 10μg RNase A (Sigma)) was added to each well containing isolated and resuspended ES cell DNA. Plates were then incubated overnight in a humidified polythene bag at 37°C. Samples were cut again by adding 10μl 1X enzyme mix (1X buffer 1mM spermidine (Sigma), 100μg/ml BSA and 10U enzyme), followed by a further incubation, in a humidified bag, for two hours.

Samples were loaded onto a 0.7% (w/v) TAE (40mM Tris-acetate, pH8.0, 1mM EDTA) agarose gel containing 0.2μg/ml ethidium bromide, alongside a 1kb ladder (Gibco) and electrophoresed at 60V until samples had travelled approximately ten centimetres (gel 300ml 20cm x 20cm). Photographs of the gels were taken on an UV transilluminator next to a ruler to allow sizing of bands. Marker lanes were removed (to avoid contamination) before gels were depurinated for eight minutes in a freshly prepared 2.5% (0.28M) HCl solution with agitation. Gels were then washed briefly in ddH_2O before denaturing for twelve minutes in fresh 0.4M NaOH. DNA from the gels was then capillary blotted overnight onto Hybond N^+ nylon membrane (Amersham) in 0.4M NaOH. Membranes were washed briefly in 2XSSC and stored wet at –20°C until hybridisation.

7.4.11 Hybridisation.

Membranes were prehybridised in an appropriate hybridisation tube (Hybaid) in 20ml of hybridisation buffer (6XSSC, 0.5%SDS, 5X Denharts solution (0.02% (w/v) Ficoll (type 400 Pharmacia), 0.02% (w/v) polyvinylpyrrolidone, 0.02% (w/v) bovine serum albumin) and 100μg/ml sheared and boiled salmon sperm DNA).

Hybridisation steps were performed in a Hybaid rotisserie oven at 65°C. DNA probes were labelled with ^{32}P-dCTP (Amersham) using the DNA Megaprime kit (Amersham), following the manufacturers instructions. Unincorporated nucleotides were removed by passing the probe through 1ml Sephadex G-50 (Pharmacia) and probe labelling quantified by scintillation counting. Probe was denatured at 100°C for three minutes before adding it directly to the prehybridisation solution already covering the membrane. Membranes were hybridised overnight at 65°C before being washed. Membranes were rinsed briefly in the tube with 2X SSC/0.5% SDS and then for five minutes in 1l of the same solution at 65°C with agitation. Wash solution was replaced with 1l 1X SSC/0.5% SDS and agitated at 65°C for twenty-five minutes. This solution was then replaced with 1l of 0.5X SSC/0.5% SDS and agitated for 15 minutes at 65°C. Membranes were then sealed in polythene bags and exposed overnight to a phosphorimaging screen (Molecular Dynamics). Screens were then processed using Molecular Dynamics Storm phospho-imaging screens and reader with Image Quant V software.

7.4.12 PCR screening of targeted clones.

ES cell clones that were deemed to be correctly targeted by southern blotting using a 3' external probe were then subjected to PCR analysis to determine the presence of the *loxP* site distant from the selection cassette. Oligo pairs ODB55/ODB56 and ODB55/ODB57 were used for this analysis, yielding products of 520 and 776bp respectively where the *loxP* site was present and 27bp less in each case where it was absent.

7.5 Production and initial treatments of wax embedded samples.

Staged mouse embryos were harvested and fixed overnight in 100% methanol at 4°C. Fixed embryos were wax embedded by standard methods (Kaufman, 1992) and then sectioned to 7μm. Wax sections were heated to 60°C for fifteen minutes prior to dewaxing in toluene. They were then taken to PBS through an ethanol gradient (100% 2 min x2, 90% 1min x1, 70% 1min x1, 50% 1min x1, 30% 1min x1 (dilutions in ddH_2O and finally PBS 5min x2) before immunohistochemical detection or TUNEL assay.

7.6 Immunohistochemistry

7.6.1 Enzyme linked antibody detection (bright field).

SOX2 immunohistochemistry was performed with a polyclonal antibody that had been raised in rabbit against the C terminal region of mouse SOX2 (excluding the HMG box) (Kamachi 1995) (see Avilion 2003). For control reactions the pre-immune serum was used at an equivalent concentration to the antibody.

1% H_2O_2 in H_2O was used for a fifteen-minute incubation, to block any endogenous peroxidase activity, followed by rinsing in water and two washes in PBS (5min each). Sections were placed in a humidified chamber and incubated with blocking solution (10% heat inactivated sheep serum in PBST), for 30min at room temperature.

Blocking solution was removed and replaced by a 1:500 dilution of the SOX2 antibody (or pre-immune serum) in blocking solution, and incubated overnight in a humid chamber at 4°C.

The antibody solution was removed and sections were washed twice for five minutes each in PBTX (0.02%tween-20 in PBS) before incubating with the secondary antibody. 1:200 biotinylated anti-rabbit IgG (Vector Labs) in PBTX with 5% sheep serum. Slides were then washed in PBTX (two times 5 min each) and signal detected using the Vector labs ABC kit and AEC coloured substrate for peroxidase (Sigma) as per the manufacturers recommendations. Samples were subsequently counterstained with methyl green (0.15% in H_2O) (Sigma) and coverslipped using glycerol gelatin solution (Sigma).

7.6.2 Fluorescence antibody detection.

Sections were incubated in a humidified chamber with blocking solution (3% heat inactivated sheep serum (ss) in PBST (0.1% TX100)) for at least an hour but usually overnight. Blocking solution was tapped off and replaced by primary antibody, at the appropriate dilution in blocking solution, coverslipped and incubated for one hour at room temperature in a humidified chamber. Sections were then washed three times for five minutes each in PBS, before diluted secondary antibody was applied. Sections were coverslipped and incubated with the appropriate secondary antibody for one hour at room temperature before being washed three times in PBS. Sections were mounted using Vector shield containing DAPI and antibody was visualised using epifluorescence. Samples were recorded using Deltavision™ and images analysed with Softworks™ and Adobe Photoshop™ 7.0.

Primary antibodies:

Rabbit Anti-mouse SOX2 (fb46) 1:50

Rabbit Anti-chicken SOX3 (Brunelli et al., 2003; Wilson et al., 2001) 1:100

Rabbit Anti-mouse SOX1 (fb43, fb44) 1:200 of each.

Rabbit Anti-phospho-histone H3 (Upstate) 1:100

TuJ1 anti-neuronal class III ß-tubulin (Covance) 1:1000

Secondary Antibodies:

Anti rabbitCY3 (Cambridge Bioscience) 1:500

Anti mouse Alexa488 (Molecular Probes) 1:500.

7.6.3 Simultaneous detection of more than one antibody on sections.

Where two primary antibodies, which have been raised in different species, have been used on the same sections both have been applied, at an appropriate concentration simultaneously. Similarly with secondary antibodies.

Where two primary antibodies, raised in the same species, have been used on the same sections one is processed first followed by detection and photography. Slides were placed in a microwave oven (850W) on full power for five minutes in 500ml of PBS followed by processing with the second primary antibody in the normal way. Appropriate controls were carried out to ensure no residual signal could be detected after microwave treatment. Photographs were then overlaid using Adobe Photoshop™ 7.0.

7.7 Beta-galactosidase staining of embryos.

Staged embryos were harvested into PBS on ice and fixed in 4% paraformaldehyde (PFA) solution (4%PFA in PBS), at 4°C, for 5-60 minutes, depending on their age. Fixed embryos were washed in embryo rinse (5mM EGTA, 0.01% Deoxycholate, 0.02% NP-40, 2mM $MgCl_2$, in PBS) three times for at least ten minutes each and then placed in staining solution (5mM $K_3Fe(CN)_6$, 5mM $K_4Fe(CN)_6$, 5mM EGTA, 0.01% Deoxycholate, 0.02% NP-40, 2mM $MgCl_2$ and 1mg/ml X-Gal (5-Bromo-4-chloro-3-indolyl beta-D-galactopyranoside), in PBS) at 37°C overnight. Stained embryos were post-fixed in 4%PFA solution at 4°C overnight. Embryos were either wax embedded by standard methods (Kaufman, 1992) and sectioned to 7μm or cleared in a glycerol series before photographs were taken. Some sections were lightly stained with eosin for histological clarity.

7.8 TUNEL assay for detection of apoptotic cells on wax sections.

Cell death kit was used as per manufacturers instructions. Briefly, wax sections were produced and rehydrated as described, section 7.3. Permeabilisation solution (0.1% TX100, 0.1% sodium citrate) was applied to samples for eight minutes, then slides were washed twice in PBS for two minutes each. Sections were covered in TUNEL reaction mixture and coverslipped before incubating in a humidified box, in the dark at 37°C for one hour. Slides were then rinsed in PBS mounted and visualised or processed for immunohistochemistry.

7.9 Production of 3D models from serially sectioned X-gal stained material

Samples were processed for X-gal staining, wax embeded and serial sectioned at 7-8μm. Slides were dewaxed slides and coverslipped without

counterstaining. Openlab software was used to digitally capture all sections in bright field with subsequent alignments and manipulations carried out using Adobe Photoshop 7.0. Series of digital images were processed though Velocity2 on a Silicon Graphics Octane workstation in order to produce 3D models and Quicktime animations (http://www.msi.umn.edu/software/velocity/tutorial/index.html).

Chapter 8 References

Akazawa, C., Sasai, Y., Nakanishi, S., and Kageyama, R. (1992). Molecular characterization of a rat negative regulator with a basic helix-loop-helix structure predominantly expressed in the developing nervous system. *J Biol Chem* **267,** 21879-85.

Akiyama, H., Chaboissier, M. C., Martin, J. F., Schedl, A., and de Crombrugghe, B. (2002). The transcription factor Sox9 has essential roles in successive steps of the chondrocyte differentiation pathway and is required for expression of Sox5 and Sox6. *Genes Dev* **16,** 2813-28.

Ambrosetti, D. C., Basilico, C., and Dailey, L. (1997). Synergistic activation of the fibroblast growth factor 4 enhancer by Sox2 and Oct-3 depends on protein-protein interactions facilitated by a specific spatial arrangement of factor binding sites. *Mol Cell Biol* **17,** 6321-9.

Amores, A., Force, A., Yan, Y. L., Joly, L., Amemiya, C., Fritz, A., Ho, R. K., Langeland, J., Prince, V., Wang, Y. L., Westerfield, M., Ekker, M., and Postlethwait, J. H. (1998). Zebrafish hox clusters and vertebrate genome evolution. *Science* **282,** 1711-4.

Anderson, S. A., Qiu, M., Bulfone, A., Eisenstat, D. D., Meneses, J., Pedersen, R., and Rubenstein, J. L. (1997). Mutations of the homeobox genes Dlx-1 and Dlx-2 disrupt the striatal subventricular zone and differentiation of late born striatal neurons. *Neuron* **19,** 27-37.

Ausubel, F. M. (1999). "Current Protocols in Molecular Biology." John Wiley and Sons, Inc.,

Avilion, A. A., Nicolis, S. K., Pevny, L. H., Perez, L., Vivian, N., and Lovell-Badge, R. (2003). Multipotent cell lineages in early mouse development depend on SOX2 function. *Genes Dev* **17,** 126-40.

Bell, D. M., Leung, K. K., Wheatley, S. C., Ng, L. J., Zhou, S., Ling, K. W., Sham, M. H., Koopman, P., Tam, P. P., and Cheah, K. S. (1997). SOX9 directly regulates the type-II collagen gene. *Nat Genet* **16,** 174-8.

Ben-Arie, N., Hassan, B. A., Bermingham, N. A., Malicki, D. M., Armstrong, D., Matzuk, M., Bellen, H. J., and Zoghbi, H. Y. (2000). Functional conservation of atonal and Math1 in the CNS and PNS. *Development* **127,** 1039-48.

Bi, W., Deng, J. M., Zhang, Z., Behringer, R. R., and de Crombrugghe, B. (1999). Sox9 is required for cartilage formation. *Nat Genet* **22,** 85-9.

Botquin, V., Hess, H., Fuhrmann, G., Anastassiadis, C., Gross, M. K., Vriend, G., and Scholer, H. R. (1998). New POU dimer configuration mediates antagonistic control of an osteopontin preimplantation enhancer by Oct-4 and Sox-2. *Genes Dev* **12,** 2073-90.

Bowles, J., Schepers, G., and Koopman, P. (2000). Phylogeny of the SOX family of developmental transcription factors based on sequence and structural indicators. *Dev Biol* **227,** 239-55.

Brunelli, S., Silva Casey, E., Bell, D., Harland, R., and Lovell-Badge, R. (2003). Expression of Sox3 throughout the developing central nervous system is dependent on the combined action of discrete, evolutionarily conserved regulatory elements. *Genesis* **36,** 12-24.

Cau, E., Gradwohl, G., Casarosa, S., Kageyama, R., and Guillemot, F. (2000). Hes genes regulate sequential stages of neurogenesis in the olfactory epithelium. *Development* **127,** 2323-32.

Cheung, M., Abu-Elmagd, M., Clevers, H., and Scotting, P. J. (2000). Roles of Sox4 in central nervous system development. *Brain Res Mol Brain Res* **79,** 180-91.

Cheung, M., and Briscoe, J. (2003). Neural crest development is regulated by the transcription factor Sox9. *Development* **130,** 5681-93.

Chiang, C., Litingtung, Y., Lee, E., Young, K. E., Corden, J. L., Westphal, H., and Beachy, P. A. (1996). Cyclopia and defective axial patterning in mice lacking Sonic hedgehog gene function. *Nature* **383,** 407-13.

Collignon, J. (1992). Study of a new family of genes related to the mammalian testis determining gene. (Thesis) National Academic Awards and National institutes for Medical Research, London. UK.

Collignon, J., Sockanathan, S., Hacker, A., Cohen-Tannoudji, M., Norris, D., Rastan, S., Stevanovic, M., Goodfellow, P. N., and Lovell-Badge, R. (1996). A comparison of the properties of Sox-3 with Sry and two related genes, Sox-1 and Sox-2. *Development* **122,** 509-20.

Cremazy, F., Berta, P., and Girard, F. (2001). Genome-wide analysis of Sox genes in Drosophila melanogaster. *Mech Dev* **109,** 371-5.

Crews, S. T., and Fan, C. M. (1999). Remembrance of things PAS: regulation of development by bHLH-PAS proteins. *Curr Opin Genet Dev* **9,** 580-7.

Dailey, L., and Basilico, C. (2001). Coevolution of HMG domains and homeodomains and the generation of transcriptional regulation by Sox/POU complexes. *J Cell Physiol* **186,** 315-28.

DasGupta, R., and Fuchs, E. (1999). Multiple roles for activated LEF/TCF transcription complexes during hair follicle development and differentiation. *Development* **126,** 4557-68.

Denny, P., Swift, S., Connor, F., and Ashworth, A. (1992). An SRY-related gene expressed during spermatogenesis in the mouse encodes a sequence-specific DNA-binding protein. *Embo J* **11,** 3705-12.

Di Rocco, G., Gavalas, A., Popperl, H., Krumlauf, R., Mavilio, F., and Zappavigna, V. (2001). The recruitment of SOX/OCT complexes and the differential activity of HOXA1 and HOXB1 modulate the Hoxb1 auto-regulatory enhancer function. *J Biol Chem* **276,** 20506-15.

Dickmeis, T., Mourrain, P., Saint-Etienne, L., Fischer, N., Aanstad, P., Clark, M., Strahle, U., and Rosa, F. (2001). A crucial component of the endoderm formation pathway, CASANOVA, is encoded by a novel sox-related gene. *Genes Dev* **15,** 1487-92.

Dong, S., Leung, K. K., Pelling, A. L., Lee, P. Y., Tang, A. S., Heng, H. H., Tsui, L. C., Tease, C., Fisher, G., Steel, K. P., and Cheah, K. S. (2002). Circling, deafness, and yellow coat displayed by yellow submarine (ysb) and light coat and circling (lcc) mice with mutations on chromosome 3. *Genomics* **79,** 777-84.

Downes, M., and Koopman, P. (2001). SOX18 and the transcriptional regulation of blood vessel development. *Trends Cardiovasc Med* **11,** 318-24.

Dry, F. W. (1926). The coat of the mouse (Mus musculus). *Journal Genetics* **16,** 287-340.

Estivill-Torrus, G., Pearson, H., van Heyningen, V., Price, D. J., and Rashbass, P. (2002). Pax6 is required to regulate the cell cycle and the rate of progression from symmetrical to asymmetrical division in mammalian cortical progenitors. *Development* **129,** 455-66.

Evans, M. J., and Kaufman, M. H. (1981). Establishment in culture of pluripotential cells from mouse embryos. *Nature* **292,** 154-6.

Fantes, J., Ragge, N. K., Lynch, S. A., McGill, N. I., Collin, J. R., Howard-Peebles, P. N., Hayward, C., Vivian, A. J., Williamson, K., van Heyningen, V., and FitzPatrick, D. R. (2003). Mutations in SOX2 cause anophthalmia. *Nat Genet* **33,** 461-3.

Fedorov, L. M., Haegel-Kronenberger, H., and Hirchenhain, J. (1997). A comparison of the germline potential of differently aged ES cell lines and their transfected descendants. *Transgenic Res* **6,** 223-31.

Fekete, D. M. (1999). Development of the vertebrate ear: insights from knockouts and mutants. *Trends Neurosci* **22,** 263-9.

Force, A., Lynch, M., Pickett, F. B., Amores, A., Yan, Y. L., and Postlethwait, J. (1999). Preservation of duplicate genes by complementary, degenerative mutations. *Genetics* **151,** 1531-45.

Friedrich, G., and Soriano, P. (1991). Promoter traps in embryonic stem cells: a genetic screen to identify and mutate developmental genes in mice. *Genes Dev* **5,** 1513-23.

Graham, V., Khudyakov, J., Ellis, P., and Pevny, L. (2003). SOX2 Functions to Maintain Neural Progenitor Identity. *Neuron* **39,** 749-65.

Gu, H., Zou, Y. R., and Rajewsky, K. (1993). Independent control of immunoglobulin switch recombination at individual switch regions evidenced through Cre-loxP-mediated gene targeting. *Cell* **73,** 1155-64.

Gu, Z., Steinmetz, L. M., Gu, X., Scharfe, C., Davis, R. W., and Li, W. H. (2003). Role of duplicate genes in genetic robustness against null mutations. *Nature* **421,** 63-6.

Gubbay, J., Collignon, J., Koopman, P., Capel, B., Economou, A., Munsterberg, A., Vivian, N., Goodfellow, P., and Lovell-Badge, R. (1990). A gene mapping to the sex-determining region of the mouse Y chromosome is a member of a novel family of embryonically expressed genes. *Nature* **346,** 245-50.

Han, K., and Manley, J. L. (1993). Functional domains of the Drosophila Engrailed protein. *Embo J* **12,** 2723-33.

Hardy, M. H. (1969). The differentiation of hair follicles and hairs in organ culture. *Adv. Biol. Skin* **9,** 35-60.

Hardy, M. H. (1992). The secret life of the hair follicle. *Trends Genet* **8,** 55-61.

Hartfuss, E., Galli, R., Heins, N., and Gotz, M. (2001). Characterization of CNS precursor subtypes and radial glia. *Dev Biol* **229,** 15-30.

Hendzel, M. J., Wei, Y., Mancini, M. A., Van Hooser, A., Ranalli, T., Brinkley, B. R., Bazett-Jones, D. P., and Allis, C. D. (1997). Mitosis-specific phosphorylation of histone H3 initiates primarily within pericentromeric heterochromatin during G2 and spreads in an ordered fashion coincident with mitotic chromosome condensation. *Chromosoma* **106,** 348-60.

Hojo, M., Ohtsuka, T., Hashimoto, N., Gradwohl, G., Guillemot, F., and Kageyama, R. (2000). Glial cell fate specification modulated by the bHLH gene Hes5 in mouse retina. *Development* **127,** 2515-22.

Hosking, B. M., Muscat, G. E., Koopman, P. A., Dowhan, D. H., and Dunn, T. L. (1995). Trans-activation and DNA-binding properties of the transcription factor, Sox-18. *Nucleic Acids Res* **23,** 2626-8.

Hwang, C. K., Wu, X., Wang, G., Kim, C. S., and Loh, H. H. (2003). Mouse mu opioid receptor distal promoter transcriptional regulation by SOX proteins. *J Biol Chem* **278,** 3742-50.

Ishibashi, M., Ang, S. L., Shiota, K., Nakanishi, S., Kageyama, R., and Guillemot, F. (1995). Targeted disruption of mammalian hairy and Enhancer of split homolog-1 (HES-1) leads to up-regulation of neural helix-loop-helix factors, premature neurogenesis, and severe neural tube defects. *Genes Dev* **9,** 3136-48.

Ishii, Y., Rex, M., Scotting, P. J., and Yasugi, S. (1998). Region-specific expression of chicken Sox2 in the developing gut and lung epithelium: regulation by epithelial-mesenchymal interactions. *Dev Dyn* **213,** 464-75.

Jung, H. S., Francis-West, P. H., Widelitz, R. B., Jiang, T. X., Ting-Berreth, S., Tickle, C., Wolpert, L., and Chuong, C. M. (1998). Local inhibitory action of BMPs and their relationships with activators in feather formation: implications for periodic patterning. *Dev Biol* **196,** 11-23.

Kamachi, Y., Cheah, K. S., and Kondoh, H. (1999). Mechanism of regulatory target selection by the SOX high-mobility-group domain proteins as revealed by comparison of SOX1/2/3 and SOX9. *Mol Cell Biol* **19,** 107-20.

Kamachi, Y., Sockanathan, S., Liu, Q., Breitman, M., Lovell-Badge, R., and Kondoh, H. (1995). Involvement of SOX proteins in lens-specific activation of crystallin genes. *Embo J* **14,** 3510-9.

Kamachi, Y., Uchikawa, M., Collignon, J., Lovell-Badge, R., and Kondoh, H. (1998). Involvement of Sox1, 2 and 3 in the early and subsequent molecular events of lens induction. *Development* **125,** 2521-32.

Kamachi, Y., Uchikawa, M., and Kondoh, H. (2000). Pairing SOX off: with partners in the regulation of embryonic development. *Trends Genet* **16,** 182-7.

Kanai-Azuma, M., Kanai, Y., Gad, J. M., Tajima, Y., Taya, C., Kurohmaru, M., Sanai, Y., Yonekawa, H., Yazaki, K., Tam, P. P., and Hayashi, Y. (2002). Depletion of definitive gut endoderm in Sox17-null mutant mice. *Development* **129,** 2367-79.

Kaufman, M. H. (1992). "The Atlas of Mouse Development." Academic Press. (Harcourt Brace Jovanovich, Publishers),

Kikuchi, Y., Agathon, A., Alexander, J., Thisse, C., Waldron, S., Yelon, D., Thisse, B., and Stainier, D. Y. (2001). casanova encodes a novel Sox-related protein necessary and sufficient for early endoderm formation in zebrafish. *Genes Dev* **15,** 1493-505.

Kirby, P. J., Waters, P. D., Delbridge, M., Svartman, M., Stewart, A. N., Nagai, K., and Graves, J. A. (2002). Cloning and mapping of platypus SOX2 and SOX14: insights into SOX group B evolution. *Cytogenet Genome Res* **98,** 96-100.

Kuhlbrodt, K., Herbarth, B., Sock, E., Enderich, J., Hermans-Borgmeyer, I., and Wegner, M. (1998a). Cooperative function of POU proteins and SOX proteins in glial cells. *J Biol Chem* **273,** 16050-7.

Kuhlbrodt, K., Herbarth, B., Sock, E., Hermans-Borgmeyer, I., and Wegner, M. (1998b). Sox10, a novel transcriptional modulator in glial cells. *J Neurosci* **18,** 237-50.

Kume, T., Deng, K., and Hogan, B. L. (2000). Minimal phenotype of mice homozygous for a null mutation in the forkhead/winged helix gene, Mf2. *Mol Cell Biol* **20,** 1419-25.

Kuroiwa, A., Uchikawa, M., Kamachi, Y., Kondoh, H., Nishida-Umehara, C., Masabanda, J., Griffin, D. K., and Matsuda, Y. (2002). Chromosome assignment of eight SOX family genes in chicken. *Cytogenet Genome Res* **98,** 189-93.

Lamoreux, M. L., Wakamatsu, K., and Ito, S. (2001). Interaction of major coat color gene functions in mice as studied by chemical analysis of eumelanin and pheomelanin. *Pigment Cell Res* **14,** 23-31.

Laudet, V., Stehelin, D., and Clevers, H. (1993). Ancestry and diversity of the HMG box superfamily. *Nucleic Acids Res* **21,** 2493-501.

Laumonnier, F., Ronce, N., Hamel, B. C., Thomas, P., Lespinasse, J., Raynaud, M., Paringaux, C., Van Bokhoven, H., Kalscheuer, V., Fryns, J. P., Chelly, J., Moraine, C., and Briault, S. (2002). Transcription factor SOX3 is involved in X-linked mental retardation with growth hormone deficiency. *Am J Hum Genet* **71,** 1450-5.

Lefebvre, V., Li, P., and de Crombrugghe, B. (1998). A new long form of Sox5 (L-Sox5), Sox6 and Sox9 are coexpressed in chondrogenesis and cooperatively activate the type II collagen gene. *Embo J* **17,** 5718-33.

Leonard, J. H., Cook, A. L., Van Gele, M., Boyle, G. M., Inglis, K. J., Speleman, F., and Sturm, R. A. (2002). Proneural and proneuroendocrine transcription factor expression in cutaneous mechanoreceptor (Merkel) cells and Merkel cell carcinoma. *Int J Cancer* **101,** 103-10.

Lewandoski, M. (2001). Conditional control of gene expression in the mouse. *Nat Rev Genet* **2,** 743-55.

Lewandoski, M., Wassarman, K. M., and Martin, G. R. (1997). Zp3-cre, a transgenic mouse line for the activation or inactivation of loxP-flanked target genes specifically in the female germ line. *Curr Biol* **7,** 148-51.

Li, M., Pevny, L., Lovell-Badge, R., and Smith, A. (1998). Generation of purified neural precursors from embryonic stem cells by lineage selection. *Curr Biol* **8,** 971-4.

Lyon, M. F., Phillips, R. J., and Fisher, G. (1979). Dose-response curves for radiation-induced gene mutations in mouse oocytes and their interpretation. *Mutat Res* **63,** 161-73.

Ma, Y., Certel, K., Gao, Y., Niemitz, E., Mosher, J., Mukherjee, A., Mutsuddi, M., Huseinovic, N., Crews, S. T., Johnson, W. A., and Nambu, J. R. (2000). Functional interactions between Drosophila bHLH/PAS, Sox, and POU transcription factors regulate CNS midline expression of the slit gene. *J Neurosci* **20,** 4596-605.

Malas, S., Postlethwaite, M., Ekonomou, A., Whalley, B., Nishiguchi, S., Wood, H., Meldrum, B., Constanti, A., and Episkopou, V. (2003). Sox1-deficient mice suffer from epilepsy associated with abnormal ventral forebrain development and olfactory cortex hyperexcitability. *Neuroscience* **119,** 421-32.

Marin, O., Baker, J., Puelles, L., and Rubenstein, J. L. (2002). Patterning of the basal telencephalon and hypothalamus is essential for guidance of cortical projections. *Development* **129,** 761-73.

Martin, P., and Swanson, G. J. (1993). Descriptive and experimental analysis of the epithelial remodellings that control semicircular canal formation in the developing mouse inner ear. *Dev Biol* **159,** 549-58.

Mertin, S., McDowall, S. G., and Harley, V. R. (1999). The DNA-binding specificity of SOX9 and other SOX proteins. *Nucleic Acids Res* **27,** 1359-64.

Millar, S. E. (2002). Molecular mechanisms regulating hair follicle development. *J Invest Dermatol* **118,** 216-25.

Mitsui, K., Tokuzawa, Y., Itoh, H., Segawa, K., Murakami, M., Takahashi, K., Maruyama, M., Maeda, M., and Yamanaka, S. (2003). The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. *Cell* **113,** 631-42.

Moll, I., Paus, R., and Moll, R. (1996). Merkel cells in mouse skin: intermediate filament pattern, localization, and hair cycle-dependent density. *J Invest Dermatol* **106,** 281-6.

Montero, J. A., Giron, B., Arrechedera, H., Cheng, Y. C., Scotting, P., Chimal-Monroy, J., Garcia-Porrero, J. A., and Hurle, J. M. (2002). Expression of Sox8, Sox9 and Sox10 in the developing valves and autonomic nerves of the embryonic heart. *Mech Dev* **118,** 199-202.

Morsli, H., Choo, D., Ryan, A., Johnson, R., and Wu, D. K. (1998). Development of the mouse inner ear and origin of its sensory organs. *J Neurosci* **18,** 3327-35.

Mukherjee, A., Shan, X., Mutsuddi, M., Ma, Y., and Nambu, J. R. (2000). The Drosophila sox gene, fish-hook, is required for postembryonic development. *Dev Biol* **217,** 91-106.

Muller, U. (1999). Ten years of gene targeting: targeted mouse mutants, from vector design to phenotype analysis. *Mech Dev* **82,** 3-21.

Murakami, A., Ishida, S., Thurlow, J., Revest, J. M., and Dickson, C. (2001). SOX6 binds CtBP2 to repress transcription from the Fgf-3 promoter. *Nucleic Acids Res* **29,** 3347-55.

Nagy, A., Moens, C., Ivanyi, E., Pawling, J., Gertsenstein, M., Hadjantonakis, A. K., Pirity, M., and Rossant, J. (1998). Dissecting the role of N-myc in development using a single targeting vector to generate a series of alleles. *Curr Biol* **8,** 661-4.

Nakamura, Y., Sakakibara, S., Miyata, T., Ogawa, M., Shimazaki, T., Weiss, S., Kageyama, R., and Okano, H. (2000). The bHLH gene hes1 as a repressor of the neuronal commitment of CNS stem cells. *J Neurosci* **20,** 283-93.

Ng, L. J., Wheatley, S., Muscat, G. E., Conway-Campbell, J., Bowles, J., Wright, E., Bell, D. M., Tam, P. P., Cheah, K. S., and Koopman, P. (1997). SOX9 binds DNA, activates transcription, and coexpresses with type II collagen during chondrogenesis in the mouse. *Dev Biol* **183,** 108-21.

Nishiguchi, S., Wood, H., Kondoh, H., Lovell-Badge, R., and Episkopou, V. (1998). Sox1 directly regulates the gamma-crystallin genes and is essential for lens development in mice. *Genes Dev* **12,** 776-81.

Nishimoto, M., Fukushima, A., Okuda, A., and Muramatsu, M. (1999). The gene for the embryonic stem cell coactivator UTF1 carries a regulatory element which selectively interacts with a complex composed of Oct-3/4 and Sox-2. *Mol Cell Biol* **19,** 5453-65.

Nowling, T. K., Johnson, L. R., Wiebe, M. S., and Rizzino, A. (2000). Identification of the transactivation domain of the transcription factor Sox-2 and an associated co-activator. *J Biol Chem* **275,** 3810-8.

Ohno, S. (1970). "Evolution by Gene and Genome Duplication." Springer, Berlin.

Ohtsuka, T., Ishibashi, M., Gradwohl, G., Nakanishi, S., Guillemot, F., and Kageyama, R. (1999). Hes1 and Hes5 as notch effectors in mammalian neuronal differentiation. *Embo J* **18,** 2196-207.

Olsson, J. E., Kamachi, Y., Penning, S., Muscat, G. E., Kondoh, H., and Koopman, P. (2001). Sox18 expression in blood vessels and feather buds during chicken embryogenesis. *Gene* **271,** 151-8.

Oro, A. E., and Scott, M. P. (1998). Splitting hairs: dissecting roles of signaling systems in epidermal development. *Cell* **95,** 575-8.

Oshima, H., Rochat, A., Kedzia, C., Kobayashi, K., and Barrandon, Y. (2001). Morphogenesis and renewal of hair follicles from adult multipotent stem cells. *Cell* **104,** 233-45.

Overton, P. M., Meadows, L. A., Urban, J., and Russell, S. (2002). Evidence for differential and redundant function of the Sox genes Dichaete and SoxN during CNS development in Drosophila. *Development* **129,** 4219-28.

Paratore, C., Goerich, D. E., Suter, U., Wegner, M., and Sommer, L. (2001). Survival and glial fate acquisition of neural crest cells are regulated by an interplay between the transcription factor Sox10 and extrinsic combinatorial signaling. *Development* **128,** 3949-61.

Parsons, M. (1997). A functional analysis of Sox3, an X-linked Sry related gene. University of London.

Peirano, R. I., Goerich, D. E., Riethmacher, D., and Wegner, M. (2000). Protein zero gene expression is regulated by the glial transcription factor Sox10. *Mol Cell Biol* **20,** 3198-209.

Pennisi, D., Bowles, J., Nagy, A., Muscat, G., and Koopman, P. (2000a). Mice null for sox18 are viable and display a mild coat defect. *Mol Cell Biol* **20,** 9331-6.

Pennisi, D., Gardner, J., Chambers, D., Hosking, B., Peters, J., Muscat, G., Abbott, C., and Koopman, P. (2000b). Mutations in Sox18 underlie cardiovascular and hair follicle defects in ragged mice. *Nat Genet* **24,** 434-7.

Pevny, L., and Rao, M. S. (2003). The stem-cell menagerie. *Trends Neurosci* **26,** 351-9.

Pevny, L. H., and Lovell-Badge, R. (1997). Sox genes find their feet. *Curr Opin Genet Dev* **7,** 338-44.

Pevny, L. H., Sockanathan, S., Placzek, M., and Lovell-Badge, R. (1998). A role for SOX1 in neural determination. *Development* **125,** 1967-78.

Plump, A. S., Erskine, L., Sabatier, C., Brose, K., Epstein, C. J., Goodman, C. S., Mason, C. A., and Tessier-Lavigne, M. (2002). Slit1 and Slit2 cooperate to prevent premature midline crossing of retinal axons in the mouse visual system. *Neuron* **33,** 219-32.

Reddy, S., Andl, T., Bagasra, A., Lu, M. M., Epstein, D. J., Morrisey, E. E., and Millar, S. E. (2001). Characterization of Wnt gene expression in developing and postnatal hair follicles and identification of Wnt5a as a target of Sonic hedgehog in hair follicle morphogenesis. *Mech Dev* **107,** 69-82.

Reynolds, A. J., and Jahoda, C. A. (1992). Cultured dermal papilla cells induce follicle formation and hair growth by transdifferentiation of an adult epidermis. *Development* **115,** 587-93.

Reynolds, A. J., Lawrence, C., Cserhalmi-Friedman, P. B., Christiano, A. M., and Jahoda, C. A. (1999). Trans-gender induction of hair follicles. *Nature* **402,** 33-4.

Rizzoti, K., Brunelli, S., Carmignac, D., Thomas, P. Q., Robinson, I. C., and Lovell-Badge, R. (2004). SOX3 is required during the formation of the hypothalamo-pituitary axis. *Nat Genet* **36,** 247-55.

Robertson, E. J. E. (1989). "Teratocarcinomas and Embryonic Stem Cells." Oxford University Press,

Roose, J., Korver, W., Oving, E., Wilson, A., Wagenaar, G., Markman, M., Lamers, W., and Clevers, H. (1998). High expression of the HMG box factor sox-13 in arterial walls during embryonic development. *Nucleic Acids Res* **26,** 469-76.

Roydon-Price, E., and Fisher, D. E. (2001). Sensorineural deafness and pigmentation genes: melanocytes and the Mitf transcriptional network. *Neuron* **30,** 15-8.

Sakaguchi, T., Kuroiwa, A., and Takeda, H. (2001). A novel sox gene, 226D7, acts downstream of Nodal signaling to specify endoderm precursors in zebrafish. *Mech Dev* **107,** 25-38.

Sambrook, J., Fritsch, E., Maniatis, T. (1989). "Molecular Cloning, a Laboratory Manual, 2nd Edition." Cold Spring Harbor Laboratory Press., Cold Spring Harbor, NY.

Sasai, Y. (2001). Roles of Sox factors in neural determination: conserved signaling in evolution? *Int J Dev Biol* **45,** 321-6.

Scaffidi, P., and Bianchi, M. E. (2001). Spatially precise DNA bending is an essential activity of the sox2 transcription factor. *J Biol Chem* **276,** 47296-302.

Schepers, G. E., Bullejos, M., Hosking, B. M., and Koopman, P. (2000). Cloning and characterisation of the Sry-related transcription factor gene Sox8. *Nucleic Acids Res* **28,** 1473-80.

Schepers, G. E., Teasdale, R. D., and Koopman, P. (2002). Twenty pairs of sox: extent, homology, and nomenclature of the mouse and human sox transcription factor gene families. *Dev Cell* **3,** 167-70.

Schlierf, B., Ludwig, A., Klenovsek, K., and Wegner, M. (2002). Cooperative binding of Sox10 to DNA: requirements and consequences. *Nucleic Acids Res* **30,** 5509-16.

Shen, J. H., and Ingraham, H. A. (2002). Regulation of the orphan nuclear receptor steroidogenic factor 1 by Sox proteins. *Mol Endocrinol* **16,** 529-40.

Silver, L. M. (1995). "Mouse Genetics: Concepts and Applications." Oxford University Press, New York.

Sinclair, A. H., Berta, P., Palmer, M. S., Hawkins, J. R., Griffiths, B. L., Smith, M. J., Foster, J. W., Frischauf, A. M., Lovell-Badge, R., and Goodfellow, P. N. (1990). A gene from the human sex-determining region encodes a protein with homology to a conserved DNA-binding motif. *Nature* **346,** 240-4.

Slack, J. M. W. (1991). "From Egg to Embryo." Cambridge University Press, Cambridge.

Slee, J. (1957). The morphology and development of ragged - a mutant affecting the skin and hair of the house mouse. I. Adult morphology. *J. Genetics* **55,** 570-584.

Slee, J. (1962). Developmental Morphology of the skin and hair follicles in normal and in Ragged mice. *J Embryol Exp Morphol.* **10,** 507-29.

Smithies, O., and Powers, P. A. (1986). Gene conversions and their relation to homologous chromosome pairing. *Philos Trans R Soc Lond B Biol Sci* **312,** 291-302.

Smits, P., and Lefebvre, V. (2003). Sox5 and Sox6 are required for notochord extracellular matrix sheath formation, notochord cell survival and development of the nucleus pulposus of intervertebral discs. *Development* **130,** 1135-48.

Sock, E., Schmidt, K., Hermanns-Borgmeyer, I., Bosl, M. R., and Wegner, M. (2001). Idiopathic weight reduction in mice deficient in the high-mobility-group transcription factor Sox8. *Mol Cell Biol* **21,** 6951-9.

Solomon, K. S., and Fritz, A. (2002). Concerted action of two dlx paralogs in sensory placode formation. *Development* **129,** 3127-36.

Soriano, N. S., and Russell, S. (1998). The Drosophila SOX-domain protein Dichaete is required for the development of the central nervous system midline. *Development* **125,** 3989-96.

Soullier, S., Jay, P., Poulat, F., Vanacker, J. M., Berta, P., and Laudet, V. (1999). Diversification pattern of the HMG and SOX family members during evolution. *J Mol Evol* **48,** 517-27.

Spokony, R. F., Aoki, Y., Saint-Germain, N., Magner-Fink, E., and Saint-Jeannet, J. P. (2002). The transcription factor Sox9 is required for cranial neural crest development in Xenopus. *Development* **129,** 421-32.

Steel, K. P., and Kros, C. J. (2001). A genetic approach to understanding auditory function. *Nat Genet* **27,** 143-9.

Szeder, V., Grim, M., Halata, Z., and Sieber-Blum, M. (2003). Neural crest origin of mammalian Merkel cells. *Dev Biol* **253,** 258-63.

Thomas, K. R., and Capecchi, M. R. (1987). Site-directed mutagenesis by gene targeting in mouse embryo-derived stem cells. *Cell* **51,** 503-12.

Tokuzawa, Y., Kaiho, E., Maruyama, M., Takahashi, K., Mitsui, K., Maeda, M., Niwa, H., and Yamanaka, S. (2003). Fbx15 is a novel target of Oct3/4 but is dispensable for embryonic stem cell self-renewal and mouse development. *Mol Cell Biol* **23,** 2699-708.

Tomioka, M., Nishimoto, M., Miyagi, S., Katayanagi, T., Fukui, N., Niwa, H., Muramatsu, M., and Okuda, A. (2002). Identification of Sox-2 regulatory region which is under the control of Oct-3/4-Sox-2 complex. *Nucleic Acids Res* **30,** 3202-13.

Torres, M., and Giraldez, F. (1998). The development of the vertebrate inner ear. *Mech Dev* **71,** 5-21.

Travers, A. (2000). Recognition of distorted DNA structures by HMG domains. *Curr Opin Struct Biol* **10,** 102-9.

Uchikawa, M., Ishida, Y., Takemoto, T., Kamachi, Y., and Kondoh, H. (2003). Functional analysis of chicken Sox2 enhancers highlights an array of diverse regulatory elements that are conserved in mammals. *Dev Cell* **4,** 509-19.

Uchikawa, M., Kamachi, Y., and Kondoh, H. (1999). Two distinct subgroups of Group B Sox genes for transcriptional activators and repressors: their expression during embryonic organogenesis of the chicken. *Mech Dev* **84,** 103-20.

Uwanogho, D., Rex, M., Cartwright, E. J., Pearl, G., Healy, C., Scotting, P. J., and Sharpe, P. T. (1995). Embryonic expression of the chicken Sox2, Sox3 and Sox11 genes suggests an interactive role in neuronal development. *Mech Dev* **49,** 23-36.

Wegner, M. (1999). From head to toes: the multiple facets of Sox proteins. *Nucleic Acids Res* **27,** 1409-20.

Wichterle, H., Lieberam, I., Porter, J. A., and Jessell, T. M. (2002). Directed differentiation of embryonic stem cells into motor neurons. *Cell* **110,** 385-97.

Wiebe, M. S., Nowling, T. K., and Rizzino, A. (2003). Identification of novel domains within Sox-2 and Sox-11 involved in autoinhibition of DNA binding and partnership specificity. *J Biol Chem* **278,** 17901-11.

Wilson, M., and Koopman, P. (2002). Matching SOX: partner proteins and co-factors of the SOX family of transcriptional regulators. *Curr Opin Genet Dev* **12,** 441-6.

Wilson, S. I., Rydstrom, A., Trimborn, T., Willert, K., Nusse, R., Jessell, T. M., and Edlund, T. (2001). The status of Wnt signalling regulates neural and epidermal fates in the chick embryo. *Nature* **411,** 325-30.

Wood, H. B., and Episkopou, V. (1999). Comparative expression of the mouse Sox1, Sox2 and Sox3 genes from pre-gastrulation to early somite stages. *Mech Dev* **86,** 197-201.

Wright, E. M., Snopek, B., and Koopman, P. (1993). Seven new members of the Sox gene family expressed during mouse development. *Nucleic Acids Res* **21,** 744.

Yuan, H., Corbi, N., Basilico, C., and Dailey, L. (1995). Developmental-specific activity of the FGF-4 enhancer requires the synergistic action of Sox2 and Oct-3. *Genes Dev* **9,** 2635-45.

Zappavigna, V., Falciola, L., Helmer-Citterich, M., Mavilio, F., and Bianchi, M. E. (1996). HMG1 interacts with HOX proteins and enhances their DNA binding and transcriptional activation. *Embo J* **15,** 4981-91.

Zappone, M. V., Galli, R., Catena, R., Meani, N., De Biasi, S., Mattei, E., Tiveron, C., Vescovi, A. L., Lovell-Badge, R., Ottolenghi, S., and Nicolis, S. K. (2000). Sox2 regulatory sequences direct expression of a (beta)-geo transgene to telencephalic neural stem cells and precursors of the mouse embryo,